PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/04, C12Q 1/68 — A1

(11) International Publication Number: WO 97/31011

(43) International Publication Date: 28 August 1997 (28.08.97)

(21) International Application Number: PCT/US97/02396

(22) International Filing Date: 18 February 1997 (18.02.97)

(30) Priority Data:
60/012,060 22 February 1996 (22.02.96) US

(71) Applicants *(for all designated States except US)*: THE REGENTS OF THE UNIVERSITY OF MICHIGAN [US/US]; Management Technology Office, Wolverine Tower, Room 2071, 3003 South State Street, Ann Arbor, MI 48109-1280 (US). BOARD OF TRUSTEES OPERATING MICHIGAN STATE UNIVERSITY [US/US]; East Lansing, MI 48824 (US).

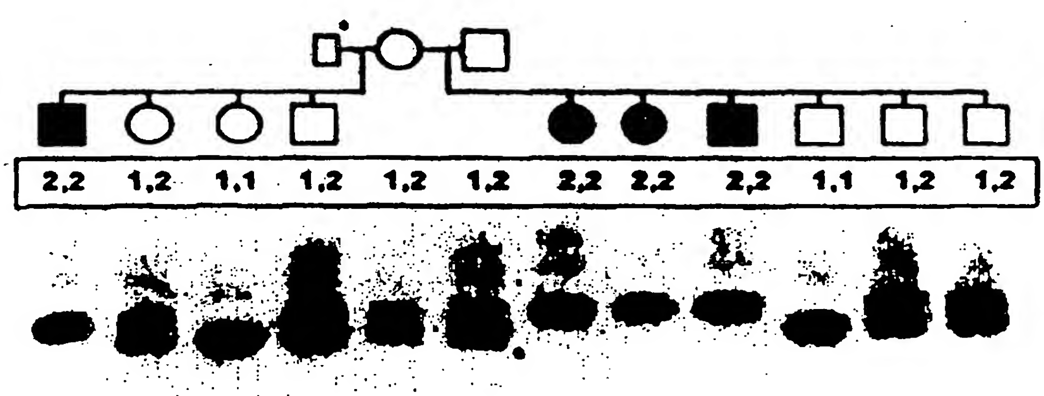
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BREWER, George, J. [US/US]; 3820 Gensley, Ann Arbor, MI 48103 (US). VENTA, Patrick, J. [US/US]; 9646 Rolling Green, Pinckney, MI 48169 (US). YUZBASIYAN-GURKAN, Vilma [US/US]; 3101 Dexter Road, Ann Arbor, MI 48103 (US).

(74) Agents: SMITH, DeAnn, F. et al.; Harness, Dickey & Pierce, P.L.C., P.O. Box 828, Bloomfield Hills, MI 48303 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MICROSATELLITE MARKERS FOR IDENTIFYING CANINE GENETIC DISEASES OR TRAITS



(57) Abstract

Microsatellite markers are provided which are useful in identifying linked markers for canine genetic diseases and traits. The microsatellite markers are derived from regions of genomic DNA which contain a repeat motif, flanked by unique sequences. The number of units contained within the repeat motif is variable, such that various different alleles are present in any given population. The microsatellite markers and their progeny are especially useful in detecting genetic diseases not phenotypically visible and identifying carriers of recessive diseases, as illustrated in the figure. In a preferred embodiment, microsatellite markers are provided which may be used to detect the canine copper toxicosis gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

MICROSATELLITE MARKERS FOR IDENTIFYING CANINE GENETIC DISEASES OR TRAITS

FIELD OF THE INVENTION

This invention relates generally to genetic markers and methods of making and using such markers, and more particularly, to a microsatellite marker that may be used to detect copper toxicosis in canines.

BACKGROUND OF THE INVENTION

Due to inbreeding and the relatively shallow gene pool, a large number of genetic diseases are present in dogs (Clark, R.D. et al., *Medical and Genetic Aspects of Purebred Dogs* (Forum Publications, Fairway, KS) (1994) and Robinson, R., *Canine Pract.* 16:29-34 (1991)). Some of these genetic diseases such as copper toxicosis in the Bedlington terrier breed, are so prevalent in a particular breed that the mutant allele frequency may be higher than that of the normal allele (Herrtage, M.E. et al., *J. Small Anim.* 28:1141-1151 (1987); and Yuzbasiyan-Gurkan, V. et al., *Genomics* 15:86-90 (1993)). Other genetic diseases cross many breeds, as exemplified by progressive retinal atrophy causing blindness (Barnett, K.C., *Adv. Vet. Sci. Comp. Med.* 20:9-67 (1976)) and hip dysplasia resulting in painful and crippling arthritis (Corley, E.A., *Small Anim. Pract.* 22:570-593 (1992)).

Canine copper toxicosis (CT) is an autosomal recessive genetic disorder of copper accumulation which results in severe liver damage. Unless specific anti-copper treatment is instituted, affected dogs die by three to seven years of age. While reported in several breeds, it is best characterized in Bedlington terriers, with the frequency of the defective gene being estimated at 50%. The disease is also prevalent in the West Highland White Terrier and Keeshond.

Currently, the only method for diagnosing affected CT dogs is by a quantitative liver copper assay in a liver biopsy sample, after one year of age. Unfortunately, heterozygous and homozygous normal animals are indistinguishable from each other by this test. In order to determine if a dog is a heterozygous carrier, test-breeding strategies must be employed which require that there be a dog of a known genotype to breed against the potential carrier. This process is very costly and results in the birth of many affected individuals. It is therefore impractical for breeders to identify breeding stock free of the gene and currently carriers of the gene are only identified after they are found to be the parents of an affected dog.

Because like CT, many of the canine genetic diseases are recessive, various methods have been investigated which would identify, on a molecular level, phenotypically normal carriers. One method that has been employed is the whole

gene subtraction method. This approach requires the sorting out of differences between DNA from those with or without the disease or trait with molecular manipulation methods. Unfortunately, this technique is somewhat impractical and requires that all variability within individuals with the trait as well as the variability within those without the trait independent of the trait, be differentiable from the one or few that are dependent on the trait. Furthermore, this method has only been demonstrated on very simple organisms such as yeast, and while this technique appears theoretically possible for higher species, it rapidly becomes impractical, as it requires many breeding studies of large numbers of affected animals.

An alternative method, the use of restriction fragment length polymorphisms (RFLP), is extremely labor intensive and expensive with respect to both characterization and analysis. Furthermore, this technique requires large quantities of DNA, generally is limited to only two alleles, and only a few loci have thus far been characterized for the canine genome. It appears that with this method, a separate genetic system must be generated for each breed of dog, and such a library may not be sufficiently variable in most situations of interest.

The randomly amplified DNA fragment length polymorphism (RAPD) approach uses random primers to amplify fragments of genomic DNA that vary from individual to individual within a species. While the primers are relatively easy to generate, the method is unreliable with minor experimental changes resulting in the resolution of different DNA band patterns. Furthermore, only a few such bands have been characterized for the canine genome.

The candidate gene method is another alternative wherein one or more candidate genes is identified based on what is known about the biochemical and clinical or other phenotypic attributes of the disease or trait and information about similar conditions in another species where a gene has been identified for a similar trait. This approach was taken in evaluating genes linked to the Wilson's disease gene in humans, a disease similar to CT. Unfortunately, the genes linked to the Wilson's disease in humans were not linked to CT in dog (Yusbasiyan-Gurkan, V. et al., *Genomics* 15:86-90 (1993)). Thus, even under the best-case scenario, the candidate gene method is merely a guess and the approach is of course, further limited by the availability of identified genes.

Because canine pedigrees for various genetic disease are abundant, with several generations and two or more affected members present in many cases, these pedigrees lend themselves to linkage studies, provided polymorphic markers

are available. Since most of the breeding is controlled, identification of linked markers would allow concerned breeders to greatly reduce the incidence of these diseases in future generations.

One type of marker that has been developed consists of simple sequence length polymorphisms (SSLPs). SSLPs arise from a varying number of repeats of a simple sequence, such as a dinucleotide repeat at a given locus, and have been reported to be frequent in most eukaryotic genomes (Tautz, D. et al., *Nucleic Acids Res.* 12:4127-4138 (1984)). Such loci, also referred to as microsatellites (Tautz, D., *EXS: DNA Fingerprinting: State of the Science* 1:21-28 (1993)), are best exemplified by those containing the $(CA)_n$ motif and are found to be highly polymorphic in many species and are being successfully used in the construction of genetic maps of the human (Weissenbach, J. et al., *Nature* 359:794-801 (1992)), mouse (Dietrich, W. et al., *Genetics* 131:423-477 (1992)), rat (Serikawa, T. et al., *Genetics* 131:701-721 (1992)) and bovine (Barendse, W. et al., *Nat. Genet.* 6:227-235 (1994)) genomes. High polymorphic information content and amenability to analysis by polymerase chain reaction (PCR) and thus to possible automation, make microsatellites excellent linkage and mapping tools.

CA microsatellites from the canine genome have been identified and their polymorphism evaluated on sets of unrelated dogs (Holmes, N.G. et al., *Anim. Genet.* 24:289-292 (1992)) or mixed bred dogs and beagles (Ostrander, E.A. et al., *Genomics.* 16:207-213 (1993)). Presently there are about 150 SSLP-type markers for the canine genome available. Unfortunately, these known markers lack the ability to detect a linked marker for any genetic trait, because of the low probability of finding a linked marker sufficiently close to a given genetic locus, to ensure detection. Many purebred dog populations have a relatively high level of inbreeding which makes it important that such markers be very polymorphic. Further, important genetic diseases occur across many dozens of breeds, requiring the markers be polymorphic in most, if not all, breeds with many different breeds having varying sets of genetic problems.

It would thus be desirable to provide a method for identifying genetic diseases and traits in canines. It would also be desirable to provide a method for identifying genetic diseases and traits in canines which has high variability and low breed specificity. It would further be desirable to provide a method which allows breeders to select and breed for certain favorable characteristics, or conversely, to avoid unfavorable diseases and traits. It would further be desirable to provide a method

which allows the detection and screening of a recessive genetic disease such as copper toxicosis, which is phenotypically undetectable in heterozygote carriers. It would further be desirable to provide a method for identifying a carrier of a genetic disease or trait and affected individuals without undergoing test-breeding experiments. It would also be desirable to provide genetic markers for the canine genome. It would further be desirable to provide a marker for the CT gene in canines.

SUMMARY OF THE INVENTION

A set of microsatellite markers are provided which are useful in identifying linked markers for canine genetic diseases and traits. In particular, five hundred and nineteen microsatellite DNA markers are provided which are highly variable within and across many breeds of dogs. The microsatellite markers are derived from regions of genomic DNA which contain a repeated motif *e.g.*, $(CA)_n$, flanked by unique sequences. The number of units contained within the repeat motif is variable, such that various different alleles are present in any given population. The unique flanking sequences may be used as polymerase chain reaction (PCR) primers which allows for the rapid amplification and characterization of each locus from a small amount of DNA. Thus, each microsatellite marker has a unique set of primers. The microsatellite markers and their progeny are especially useful in detecting genetic diseases not phenotypically visible and identifying carriers of recessive diseases. In a preferred embodiment, microsatellite markers are provided which may be used to detect the canine copper toxicosis gene.

In addition to identifying canine genetic diseases such as copper toxicosis, the microsatellite markers may also be used to create a genetic map of the canine genome, generate specific breed profiles, settle parentage disputes and identify dogs by DNA fingerprinting. Pedigrees of affected individuals, their siblings, parent and progeny can also be created. Breeders and owners can thus choose breeding stock thereby reducing and possibly eliminating the incidence of specific genetic diseases.

Additional objects, advantages, and features of the present invention will become apparent form the following description and claims taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The various advantages of the present invention will become apparent to one skilled in the art by reading the following specification and by referencing the following drawings in which:

Figure 1A is a bar graph showing the average and standard deviation of heterozygosity percentages across loci within a breed;

Figure 1B is a bar graph showing the average and standard deviation of heterozygosity percentages across breeds within a locus;

Figures 2A-2D are photographs of gels showing marker locus D02011 in various breeds; and

Figure 3 is a photograph of a gel showing segregation of alleles at the C04107 locus in a Bedlington terrier pedigree.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Five hundred and nineteen microsatellite markers from specific gene loci are provided which are highly variable within and across many breeds of dogs. The microsatellite markers of the present invention comprise a repeat motif *e.g.*, $(CA)_n$, found in the canine genomic DNA, flanked by unique sequences. The unique sequences (also referred to herein as primer pairs) may be used as PCR primers, allowing the rapid amplification and thus detection of the sequence of interest in a small DNA sample. Table 2A sets forth the microsatellite markers of the present invention. The microsatellite markers and their progeny are especially useful in detecting genetic diseases not phenotypically visible and identifying carriers of recessive diseases.

In a preferred embodiment, microsatellite markers are provided which may be used to detect a carrier of the canine copper toxicosis gene. As further set forth in Specific Example II below, marker locus C04107 may be used to predict the inheritance of alleles at the copper toxicosis locus. C04107 has also been used to locate two other marker loci C04107B and C04107C, which either singly, or as a group, may also be used to detect the copper toxicosis gene.

The method of the present invention is useful for identifying disease free individuals (homozygous normal), carriers (heterozygous) and affected individuals (homozygous affected) at any stage of development. While a single marker may fail to provide the required information in any particular pedigree, a series of progeny

markers will, and thus such a family of progeny markers derived from the linked markers set forth herein, are also included in the invention.

SPECIFIC EXAMPLE I

Materials and Methods

Isolation and Characterization of Microsatellite Loci. Established protocols were used for the cloning and screening procedures as described (Sambrook, J. et al., *Molecular Cloning. A Laboratory Manual* (2nd ed. Cold Springs Harbor: Cold Springs Harbor Laboratory Press) (1992)). Genomic DNA was isolated from a peripheral blood sample from a Labrador retriever and partially digested with *Bam* HI. Size selected fragments purified from agarose gels using QIAEX beads (Qiagen Corp., Chatsworth, CA) were cloned into the phagemid vector pBS (Stratagene, La Jolla, CA) to construct a library of average insert size of 600 bps and propagated in the host XL-1 blue. The library was plated at low density (about 500 colonies/plate) without amplification. Duplicate nitrocellulose colony lifts were prepared, denatured and hybridized with $(CA)_{16}$ oligomer, labeled with ^{32}P dCTP using terminal transferase. Positive colonies were picked with a sterile pipette tip and lysed in 50 μl of a solution consisting of 1% Triton X 100, 20 mM Tris and 2 mM EDTA. Using primers complementary to the T3 and T7 promoter sequences which flank the cloning site, the inserts were amplified from 1-2 μl of the colony lysate in polymerase chain reactions for 30 cycles of 94, 55 and 72°C at 1, 2 and 3 min., respectively after an initial denaturation at 94°C for 4 min. The standard buffer, nucleotide and primer concentrations were 50 mM Tris-HCl (pH 8.3 at 25°C), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM dNTPs and 40 pmoles of each primer in 100 μl reactions. PCR reactions were carried out on either a Perkin-Elmer Cetus (Perkin Elmer, Corp. Norwalk, CT) or an MJR PTC-100 thermocycler (MJ Research, Watertown, MA). To carry out secondary screenings of the clones, aliquots of the amplification products were run on 1.5% agarose TBE gels (90 mM Tris, pH 8.3, 90 mM boric acid, 2 mM EDTA). Southern blot analysis was carried out on the gels after transfer to Gene-Screen Plus membranes (NEN, Boston, MA) using the alkaline transfer protocol. The membranes were probed with $(CA)_{16}$ oligomers, 3' end-labeled with digoxigenin-dUTP using terminal transferase. A chemiluminescence detection system based on Lumi-Phos 530 as a substrate was used to detect positive hybridization signals following the recommendations included in a commercial kit, Genius (Boehringer Mannheim Corp., Indianapolis, IN). The membranes were washed to a final stringency of 0.1 X SSC (1 X SSC = 15 mM sodium chloride, 1.5

mM sodium citrate) at 65°C. The blots were than processed for immunological detection as described by the manufacturer. Once a clone was confirmed to be positive, the corresponding amplification product was then purified using QIAEX beads (Qiagen Corp., Chatsworth, CA) after electrophoresis on TAE gels (40 mM Tris acetate, pH 8.3, 2 mM EDTA) and directly sequenced using cycle sequencing (Delta Taq 2.0 Cycle Sequencing Kit, United States Biochemical Corp., Cleveland, OH). The sequencing reactions were carried out according to the manufacturer's instructions with the slight modification that T3 and T7 primers labeled at their 5′ end with ^{33}P ATP (NEN, Boston, MA) using T4 polynucleotide kinase were used as sequencing primers. Sequencing products were analyzed by electrophoresis on 6% polyacrylamide gels containing 8M urea. The gels were dried and exposed to X-OMAT X-ray film (Eastman Kodak, Rochester, NY) for 1-2 days and developed. Primers flanking the repeat motif in each insert were selected to minimize hetero- and homedimerization; occasionally, the computer program Oligo (National Biosciences, Plymouth, MN) was used to help in the primer selection. The primers were synthesized by the Michigan State University Macromolecular Structure Facility.

Dog DNA Panel. To check the usefulness of microsatellite markers within and across different breeds of dogs, a dog DNA panel was established. The breeds to be included in the panel were chosen with consideration given to the diversity in origin and function of breeds that exist. Table I presents various characteristics of the breeds chosen for the dog panel (Alderton, D., *The Eyewitness Handbook of Dogs* (New York: Dorling Kindersley) (1993); American Kennel Club, *The Complete Dog Book* (17th ed. New York: Howell Book House) (1985); Clark, R.D., *Medical and Genetic Aspects of Purebred Dogs* (Forum Publications, Fairway, KS (1994), Walkowitz, et al., *Successfuly Dog Breeding* (2nd ed., New York, Howel Book House) (1994); and Lee, M.P., *The Official Book of the Scottish Terrier* (Neptune City, T.F.H. Publications p. 158) (1994)). Five to ten individual dogs from each breed were selected for inclusion in the panel. Pedigrees were investigated to ensure that only dogs that had no common ancestors through four generations were included for independent representation of alleles. Ten, apparently unrelated, mixed bred dogs were also sampled. DNA was isolated from peripheral blood as previously described (Sambrook, J et al., *Molecular Cloning. A Laboratory Manual.* (2nd ed., Cold Springs Harbor, Cold Springs Harbor Laboratory Press) (1989)).

Table 1

Various Characteristics of Breeds in Dog DNA Panel

Breed	Country of Origin	Current Classification	Date of Origin	Height Range (cm)	Weight Range (kg)	Litter size
Cocker Spaniel	Great Britain	Sporting Dog	1800s	36-38	11-13	5
Labrador Retriever	Canada	Sporting Dog	1800s	51-57	25-34	7
Pointer	Great Britain	Sporting Dog	1600s	61-69	20-30	6-16
German Shepherd Dog	Germany	Herding Dog	1800s	57-62	34-43	8-10
Shetland Sheepdog	Great Britain	Herding Dog	1700s	35-37	6-7	4-6
Beagle	Great Britain	Hound Dog	1300s	33-41	8-14	5-6
Greyhound	Great Britain	Hound Dog	3000 BC	69-76	27-32	10-15
Scottish Terrier	Great Britain	Terrier	1800s	25-28	8.5-10.5	3-6
Doberman Pinscher	Germany	Working Dog	1800s	65-69	30-40	8
Siberian Husky	Siberia	Working Dog	1800s	59	16-27	3-7

Analysis of Microsatellite Variability. Amplification of the correct target was verified by comparing the product obtained from genomic DNA to that obtained from the reference clone. The variability at each locus was tested by amplification of DNA from the dog panel. PCR conditions were 35 cycles of 94°C, optimal annealing temperature (50-60°C) and 72°C at 1, 1, and 2 min., respectively after an initial denaturation at 94°C for 4 min. in the standard PCR buffer conditions described above. 100 ng of genomic DNA was used as template in each reaction. 10 μl of the PCR products were analyzed by vertical electrophoresis using a modification of a SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) protocol (Laemmli, U.K., *Nature* 227:680-685 (1970)) as described previously (Tas, S., *Anal. Biochem.* 188:33-37 (1992)). An HSI SE600 vertical slab gel electrophoresis system (Hoeffer Scientific Instruments, San Francisco, CA) connected to a cooling unit was used. The gels were poured between 16 x 16 cm. plates using l mm spacers. 1.5% acrylamide stacking gels of 2-3 cm were used on top of 12.5% acrylamide separating gels with 30:0.8 acrylamide to bis-acrylamide ratio. The gels were run at 40 mA through the stacking gel and than at 70 mA thorough the separating gel until the bromophenol blue dye reached the end of the plates, for approximately 4 hours. The amplification products were visualized after silver staining with the Silver Staining Kit (Bio-Rad Laboratories, Richmond, CA). This procedure resolved differences greater than or equal to 4 bps in the size of amplification products in the 75-250 bp range.

Results

Screening 110 plates resulted in the isolation of 1064 independent clones that were confirmed to be positive on secondary screening. Using 600 bps as the average insert size and 500 as the average colony number per plate, it was calculated that 1064 positives reflected an estimated incidence of one CA repeat clone every 31 kilobases in the dog genome.

The first 14 CA repeat loci for which primers were designed are presented in Table 2 together with the optimal annealing temperatures.

Table 2

	Marker Locus	Primer Pair	Repeat Motif in Reference Clone	Product Size (bp)	Annealling Temperature °C
1	D00101	ACTCTTCTCCATCTCCCTCTGC TCGTTGGGGTTAAAGCTCTGACC	(CA) 9	150	65
2	D00401	TGCCCTCACCAGGTGTATAGA GTGTGAATATGATGTGTCTGAAAA	(CA) 22	90	58
3	D01205	AGCATGATGCCCTTCAAGGTC GGATCTTTACCCGCATGTTCC	(GT) 16	201	58
4	D01902	CCTACTAAAATACAGAAACG AACTGTTAGAACTTAGACATGC	(CA) 18	129	55
5	D02001	GTTCTCATAGAAGGAAGTAGGAGC ATATTCTCTTAGGTTAGACAGCAGG	(CA) 20	270	67
6	D02005	TCTAAATATGATTATGTATGCGT CACTTTATAACAACATATTCAAAT	(CA) 13	119	55
7	D02011	GGTCACCAAGCTAAGAATGTTGC GATCTCTCTTGCTATTGCTC	(TA) 7 (CA) 13	238	55
8	D02012	CTGAGATGTGTCAAAAGTCCTTTCG TTGCCTACAAGATCCCTACATGCC	(CA) 15	171	60
9	D02202	TTAAGCAGAAGCTCCGCTGC AATTTTGGTGCCCACTATGGAAGCC	(CA) 12	91	60
10	D03709	ACATTTCTGAGTGGCATGGCT ACTCCCAAATCTTCACAAAGGAA	(CA) 9	86	58
11	D03805	GTCAACAGCTTAGAAGTCACCA ACTATTATGCTGTAGGGGTGCAA	(CA) 12	90	58
12	D03908	TACACCTGACACTTGTATCC GTGCTTGTTAGTCCATGACC	(CA) 13	94	58
13	D04403	CTATTGATTTTTCCAAAGC GTCTTTCATGTTTTCATATACTC	(CA) 15	130	50
14	D04702	GTCTTCCAAGTGGTAAGAGCCTACC ATCCTCCTCTACCCTCAGAGCC	(CA) 12	112	60

The complete set of microsatellite markers is set forth in Table 2A below. These markers were identified and the primers designed as described above.

Table 2A

Marker Locus	Sns Sequence	Asn sequence	PCR Product (bps)	Motif
C00103	CTACTCTTGTATTCCATCAAT	ATTTTCCCCATTCTCACTGGT	242	(GT)21
C00104	TGACATAAGCTGTGAGAAGAC	ATTGAAACTGATAGAGAAGAG	140	(GT)9
C00111	TACGGAGCCCACACTACTGA	TCCAAGGGAAGTCATAGAAC	226	(GT)11
C00111	AGCTTCCAGGTCTGGTTTTCCAAG	TATCCCAGAGCTTAGAGCCTGGCA	174	(GT)11
C00113	TTTTTGATGGCTGAATAATA	GAATGGATAAAGAAGATGTG	82	(GT)14
C00114	CTGCTTCTCCCTCTGCCTATGT	CTACCACAGCCAATGTTGATTGA	140	(GT)12
C00203	AGGGTGCCTAACTGACTGAGCC	TTTCAAAATGGGCTTTCCTTT	162	(AC)17
C00203	AGGGTGCCTAACTGACTGAGCC	TTTCAAAATGGGCTTTCCTTT	162	(AC)17
C00215	TGCCCCTTAAAGATTTTATTT	CCTGCATCGAACCTGCTTCT	127	(CA)10ACT(AC)12(AG)4
C00217	TCCTGCATGGAGCCTGCTTCT	TGTGTATTCAGATGTGCTACTTGGT	181	T11A2G(AT4)(AT3)2(AT2)__(AC)10...(GA)16
C00304	GCACCACTTGTAACCCTTGAAC	TCGCATAGGATGATGAATAATA	181	(CA)4TA(CA)12
C00403	ATGGAGCCTACTTCTCCCTC	GACTTGCTGTATTGGTTACACT	123	(TG)11
C00412	ATCAGTCCATTCTGATTGGCTATC	GAAAATGGCAGTTGTACCTGAATCT	209	(TG)13(TA)4
C00501	ATCACATCCAAATCAAGACTAT	TGTCCTATGCCTGTCCTATTAT	172	(AC)15
C00502	TGACTTTACCTTACTTCACCTT	AGGGCAACTTGGTTACAGATTA	109	(CA)3T(AC)2C2(CA)6
C00505	CAGAGCCTTCAGATAACAGTA	ATTATTCTTTCCCTTTTCTAC	230	(GT)9T(TG)4(TA)4(TG)7
C00506	CATATCCATCCTCCTAAACTTTC	AGTGCCTAAAACTAACAGAACTG	173	(GT)2A(GT)9
C00602	CCAGGAAGTTATGATTCTAAATGT	GAGCTTGCTTCTCCCTCTGCC	214	(AC)7(AG)8
C00603	CTTTTCCTATTGTCACAAATG	ACAGATGAATGAATACAGTTG	107	(TG)12
C00607	AGTCCCACATCGGGCTCTCT	TGCTGGTTTCTCTCTTGTGTCTTAT	169	(CA)9TA(CA)4
C00613	GTGGAGCCTGCTTCTCCCTCTG	CTTCCAAGTGCAAACACATAGC	191	(GT)7(A3T)n
C00802	TACCTGAGTCAGTTTACCTAGCA	GTTTCTACAGTCAACCAGATG	185	(GT)19
C00803	TAAGAGTTATGCCACTTGACC	CCAGGGAAGAGACCAGTATATGA	100	(GT)12
C00901	TAAAGGTCCATTGATAGAGGA	TGATCCCAGGAGTTCATTCTT	105	(AC)12
C00902	GAGCCTGCTTCTCCCTCTG	TGTTTCTTCAATGACCTTTCAG	175	(CA)14
C01001	ATGGGCTCCAAGAATAGCA	ACCAGAAACTTCATTGTCTCC	219	(GA)12
C01003	GAAGTAAATCAACAAACAATCA	GAAGCAAAAGTATAAGAGCTGTG	87	(AC)11
C01201	ATTCTTTCTATGGCTAGGCAGT	TGAGTTTCTCCCTCTTTCTCT	150	(GT)6A(TG)5A(TG)3
C01207	AGACCACTCTGCTCCCTCTT	TGCCTTGAAATGAACAATGA	84	(GT)15
C01212	AGGTGTTCTCACTCCTCATA	CTCCCTCTGCCTGTGTCTCT	115	(CA)10
C01304	CTGAGCAAGACCCATACCACTT	CCTCCCCAGAACAATCTATTTC	180	(TG)7TA(TG)4
C01305	GCATGAGATAAGACACCACCTGTT	TTCATTTCCTGCCTCCTGTG	136	(GT)9
C01403	GAGGCTGACAACTGTTTGCTA	GGAGATAAATGATGAGAACTCA	284	(AT)2T(AT)7CA(GA)4__(CA)7(GA)2(CA)2
C01406	GATTTTATTCATTTATCCATGAC	CTCCCTCTGCCTATGTCTCTG	107	(CA)16(GA)16
C01406	TGGTGAAAGTAACTAAGAACA	TCCCTCTGCCTATGTCTCTG	150	(CA)16(GA)17
C01409	GTTCTTCCCCAATGGTATTTA	TTGCATAAGAGCCAGCAAACT	246	(CA)6A2(CA)3
C01505	TCTGCCTATGTCTCTGCCTGT	ATAAGATACACGAACCATTAGCC	109	(GT)13
C01601	CCTGCATGGAGCCTGTTTCTC	CATTTCTGGAAGACATACTGGTA	145	(GT)7
C01606	ATGCTGTTGATTACACAGACC	ATCACTTCCTGGTATTCACAC	109	(GT)19
C01801	TCTGATTTTCACCCTTAGAAC	GCAGTTTTCCTGTCTCTCTT	144	(TG)10(GT)9
C01802	ATGCAAGTTCTAAAACCATACTG	TAGTGAAGACAGGATTGTGTTG	137	(TG)19
C01908	ATCAAGTCCCACATCAGCAGCCT	AGTGGTATGAGGGGCATAAGGAA	189	(GT)10
C02005	GAGTAAAGAAAGAGTTTGAACAAT	AGTTGGAGAAATGAGCACTTA	146	(GT)10
C02122	ATGTCAGGCTCCCTGCATGG	GTTAAATGTAAGATGTC[illegible]CAGCCTTT	149	(CT)4GT(CT)6(GT)6(CT)3
C02401	CCAGACCCAATGACATCTCC	ACCCAGGTGCCCTCTTATCC	236	(GT)18
C02509	TGGCCTAAACACCTCTGACAT	TGGGATACAAAGTAAATGGAAC	189	(CA)18
C02511	GACATGATTACCACATTCATC	GTACAACTGAAGAGACTGACC	97	(GT)16
C02601	CTCCCTCTGCCTGTGTCTCT	TGTTAGTCTTAGCCATTCTGA	144	(GT)8(CT)3__(CA)12
C02604	CTCACCCAGAGGATGCTTTGAA	TTAACCTGAGAACATGGCACAA	190	(CA)17
C02608	AGGGAGCAGGTTTGTGGTTG	TACTTCTGGTCCAACATTTCC	110	(GT)19
C02705	GAGTGATTCTCATTCAAAAAGGGA	TCAAGGGCACTTTCTACTGTGTA	116	(GT)10
C02709	CTCTGCCTACGTCTCTGCC	CACCAGTATGCTGATATAATTCT	142	(CA)18
C02711	TCTCATTCAAAAAGGGAGATGC	TTTCAAGGGCACTTTCTACTG	109	(GT)10

Table 2A (cont.)

C02712	GCTTGGATGCTATTGGCTCAA	CAATGACTTGGGAAACTACATTC	156	(GT)22
C02802	CCCTGCATAGAGCCTGCTTCT	AGCTTTTGCTTATTATATGCTTG	186	(GT)6(CT)2CA(TG)6(TC)2(AnT)
C02805	GACAAGAACAGGTATGAGAGC	TGTTGAGTGTAAGATTCAAAGC	118	(CA)12
C02806	TCCCTCCTCCTGTGTCTCT	CTACACCTGTGAAACTACCA	159	(GT)11GAG(A3T4)(CT3)A6(TA2)3(TA3)
C02903	CCTACATGGAACCTGCTCTTC	TGTCTTTCCCTCAACAAGATG	167	(TC)4TG(TC)6
C02911	ATCATGGGAGAGGGTGGTAT	GGGTAGATAAAGACCTGTAAG	122	(CA)16
C03001	TTCAGAGTTAATGATGCTTAGG	GAGATTCTCTCCCTGTACCAC	153	(GT)7(GA)17
C03102	ACTTGTGTTACCCCTTTTACC	CCTGCCTTTATGGAGTTTACA	108	(CA)5TA(CA)15
C03104	TCCCTCTGCCTGTGTCTCTAC	ATCAATGAAACAAAAGGAACAGTA	147	(GT)19
C03109	CCTGCATGGAGCCTGCTTCTC	CACACCAATTAAACAATAGACATT	185	(GT)16
C03301	CCATTCCCATAGAGAGGAA	ACCTAGCCAGGACTGGAAAG	118	(CA)7TA (CA)11
C03302	TGAGTATTATGACCTGGAGGGT	TCAGTAGGTTGTGTCTAGCCT	97	(GT)11C(TG)5
C03302	TCTCAATGATACAAGAACTTCAC	TCCAGTCACCCTCCAAGATGT	185	(AT)11(TA)8(CA)16
C03304	ATTGGCATCATTCCACTGGTCA	TGGAGGCAGCTTAAATCTCAACA	95	(AC)16
C03308	TGATAAGAGTGTGAACAGAGAAGA	CTAGGAGATTGTACAGGTGCT	275	(GA)~20
C03401	GGTCATCTTTATACCATCAATTAG	CTTTAATGCTGGCAGATGCTAT	104	(CA)10
C03404	CAATTCTCTCTATGCCTCTTTGT	TCTTCTTGATTCACAGCCAATCT	171	(CT)4T(CT)2GT(CT)10(CA)18
C03501	TCGGAGATGGAAACTTTTGTAAGAG	TCTAGTGGACTGTTCTGAATTTG	106	(GT)21
C03507	ATCTCGTAATTTCCCATAATACTTA	ATCAAGTCCCACATCAGACTCC	161	(GA)2(CA)5TA(CA)6(GA)6
C03508	TACTCCAATGGCAACAGTTTA	CCTTAGACCATCTACCTCTTTTC	110	(CA)5G(CA)17
C03509	CATTCTGCTCATCTCCATAAG	GGCACAACTAACTCATTTCTAT	188	(CA)15
C03510	CCTGCATGGAGCCTGCTTCTC	TGGCTATTTATGGAGCATCTCTT	156	(GT)19
C03512	GAGCCTGCTTCTCCCTCTG	GAGACCATAATTCACAATTCTTC	113	(TC)12ATGA2T(A3)T3___An
C03601	AGCCTGCTTCTCCCTCTGTC	TGTTGCTTACCCTTCTGTTAGA	151	(CT)3(GT)10(CT)2
C03607	AGTTCCATCCACATCGTTGCA	AGAAAGAGCCTAGATGCCCAT	141	(GT)18
C03810	TGCTTCTCCCTCTGCCTGT	GGCTGTAAGACGCAGATTTCT	134	(AC)17
C03814	ACATTGGGTTCCTGCATGGAG	GGCAGTTTGGTGATGTCTATCAA	237	(TG)19
C03815	GTGCATGGAGCCTGCTTCT	AGCTTAGCACCCTGCATGGA	161	(CT)6___(TA3)2(TA5)(T2A4)(TA3)4
C03907	TAGTGCTCATGGAGCCTTTCA	TATGCTGATTCCACCTACCTC	83	(GT)13
C03909	TCAAATCAACTCGTGTTTCTGT	GGATCTGATAATCCACTTTAGA	71	(TG)8
C03913	GAAGGGACAGAGAAAGAAATGAC	TGTAAGGGCTGTTACCTCTAATC	333	(TC)13(AC)12
C04003	GGGTCTCCTTATCACACTG	AGCAACACTTGACATTATTT	135	(CA)12
C04007	ACCAAATGAGCCACTTAGGT	CCTCTGCCCTTTCCTCTATG	109	(CA)11
C04103	AATGCTGTGGAAGGTGAATGATA	ATGGAGCCTGCTTCTCCCTCTG	224	(CA)8(GA)4
C04107	TCAGCAACTATACATTTAAGAGCA	CTGTCCCATCTAAAGGATAGG	160	((GT)6GA(GT)11
C04107B	ATCGAGTCCCACATCCTTG	CATTTACTGGTTTGTCAGTTAGG	120	(AG)11
C04107C	TGGGAGATGAAAAGTATCCTC	CCTGTGCCTCAAGATAGATG	250	(CA)18
C04201	GAGTTCCTTCTTCCGCATCTAG	ACTATTCAGAAAGCAGTACAACCT	120	(GT)6A2(GT)14
C04208	ATCCTAGTTAGGCATGTGCTT	GGTAAATTACAGCAGGTGAT	205	(GA)2(AC)11
C04302	TGGTTATTACTGAGCAGACATC	GCTTTTGTTTCCTTCAAATAC	168	(GT)21
C04601A	AGAACCTATCCAGCTATTATAGTG	CTCTCAGATATGACCAACCTA	214	(TG)18
C04601B	ATATACTTTCACTCTCCATGCAA	AGAAGAGGAGTCTTTGGATG	139	(TG)18
C04704	CAGTTGCTAAGAGGTAGGTC	GTAAATGATTACCATAATAAGGT	114	(CA)13
C04716	TTCTCCCTCTGCCTATGTCT	AGCACCCTGGTACTGTTTCT	133	(CT)3(GT)9(ACT)(ATC)A(TA3)2(TA8)(T3A12)
C04802	TTACCAAGCTAAGCCTGGCA	TGGAACCATCACTGAAGGGA	150	(C6A)(C6T)(AC)20
C04802	AGACCACCGAATGGATGGAGT	TGGAGTAAGTAGCAATCCTCT	144	(AC)17
C04805	CTTTGGTCTCTGGTGGCAATAG	TGGACTTGTGATACACCGCACT	207	(CA)17
C04806	GCCTCACTCATCATCATTTTC	GAACAAGAGATTCATATTTGCTATCA	180	(TG)18
C04903	ACTGCAAATAACCTGTAGAGTGCT	ACCAATCACCATTCCCTCATTC	157	(AC)16
C04904	AAGACTTCACCACTCACAGTCA	CTGGCTCAGTGTGTATGAATG	143	((CA)6T(CA)11
C05101	CTCTTAACCGACCTTGACACC	AGAACTTGCTTATGAAGTCATGT	208	(AC)15
C05102	AAGCTGTGATGTGGCTCTCAAC	CAATGGGCAGAAACAATGAGGA	171	(AC)20
C05103	ATTGGCATTTATCTTCATTGT	AAGAGGAAAGAATCTGTGAACT	196	(GT)16T(GT)2A(TG)5
C05110	TGGAGCCTGCTTTTCCCTCT	ACCCTGAGACCATGAGCTAAG	185	(CT)5(GT)8

Table 2A (cont.)

C05112	GTACTAACTCCTTGCATTTCATC	GGCACCAAGTGTTTTCATGTAAT	138	(CA)2CG(CA)9
C05201	CTGCTTGAACACTGCCATC	GGCATGGAGCCTGCTTCTC	167	(CA)18
C05204	GAGCCTGCTTCTCCCTCT	TACCTGTCACCATACATAGT	164	(CT)2(GT)14
C05205	ATCACGACCCTGAACCTAAG	CCTGCTTCTCCCTCTGCCT	224	(AT3)10.....(AT)3....(AC)10
C05206	TGACCTTGGGAAGCTGGAG	CCATCAGTGGTGTTATCTGTA	151	(GA)2G(GT)14
C05302	GAGCCTGCTTCTCCCTCTG	CCAGGATTTGGAAGGTTCT	178	(GT)15
C05303	ATCAAAGTGACACATCATATT	TGAAAGGACGCTGAATTGG	132	(AC)18
C05305	TATTGCATCCTGCTTCCAGA	CAGCCACGTTGGCCCTTCT	105	(GT)14
C05306	ACAATAGCCTAGATATGGAAGCA	GCTGCAAATAGCAAGAATTCAT	148	(TA)3(CA)13
C05307	TGAAGTAGTAGCCTAACTGACA	TAATCCTAATCCACTCTAATGGT	300	(AC)15
C05401	CGGTGCATGGAGCCTGCTTC	CTGAACCATCCAGATGTCCAGA	152	(GT)13
C05403	GGTGCATGGAGCCTGCTTCT	CACCTACCTCCCCTTCTGCAA	141	(CT)3(TG)10
C05404	CTGTATGGAGCCTGCTTCTC	CCTTGAAGGATATTGTGTCC	138	(CT)3(GT)13(CT)2
C05405	CTAAACCACTGAGCCACCTG	ATGTGTAACAGAAGCCACTAA	263	(GA)2(CA)6TG(CA)7
C05406	CAGGGATCTTGCTTTTAGCAT	ATTGATGTTTTGTCAGATTC	280	(TG)3TA(TG)7
C05407	ATTATTACTGGTGGCTTATTTAGA	TCATGGGTCTAAGTGTTTGGA	101	(CA)8
C05409	CGGTGCATGGAGCCTGCTTCT	GGGAGATAGACAATCACCAAAT	231	(CT)15(GT)7(CT)2
C05410	TTTCAGTCCAGCCAAATGAAC	CCTGGGATGGAGCCTGCTTCT	183	(CA)8
C05414	GAGTCCCACATCAGGCTCC	GCTGTTTACACAAAACATAGAAG	150	(GT)11
C05415	GCCACCCAGGGATCTTAAAT	CCATTACCTCACATGGTTACTT	73	(AC)7
C05503	TACCACTCTGCTTGGACAT	ACTAATTCCAATGTACTGTTAC	163	(AC)9
C05504	GTCCACTTCCAATTGCCGTT	AAGTACAGGAATTCTGTTATGAG	234	(CA)2G(AC)8
C05505	AATCTCTCAAATCTCCTCCAT	CTCTGATTCCTCTAGTTTCTTTCCT	243	(TG)11T3(GT)4
C05506	CACATGGGCCAATTCCTATAA	GTATTGGTCAGGATTCTCCAG	136	(CT)17(AC)7C(CA)10
C05509	TGTCGGTAGCATAGCATAGAA	CCTCAGTTTTACATGAACTCA	78	(CA)14
C05601	CTGCTTAGAGTGCTGTACCAC	CTCAGCTCCTGGACACTTCCT	168	(AC)19T(CA)4
C05602	TCTAGAGGATCACATGCAA	CTTCTGGACTCCTGCCTTCC	105	(TG)15
C05604	CAGATGTTCAGAATGATTTAATAG	ACCTGATATGTGGCATGTTGT	227	(AT)4(GT)7
C05606	TATAGTAGGATTCTTGTGGTTG	ATCGAGTCTCACATCGGGCTC	194	(AC)23
C06105	AATAATGAAAACAGCCAACTT	ATCATAATGATTGAATGAGAT	98	(GT)12
C06106	AATAATGAAAACAGCCAACTT	TTATTTAACCCACTGAGCTACC	151	(GT)12
C06114	CTCCCTCTGCCTGTGTCTCTG	GGGCTCTTCCTTTGTATCTTT	140	(GT)14
C06201	TCTCCTTCTGCTACTTCTCC	TAGTGGTGGGGTTGAAAGAG	138	(A3T)11
C06204	GGCTGCCCCTCACACATATT	ATAACATCTGGATTGGGTCTA	105	(CA)10TA(CA)8
C06213	CTGATATAGGTAAGTTGCATTTTG	CTGGAGCCTTTTAAGGTCATT	177	(GT)14
C06216	ACTCTCTTCCTGGCTTGTAGATG	TAGCACTCTCCCCTTCCCCTTA	167	(GT)15
C06404	ATCAACCACACGCTCCTTCTT	TTGGGGGAGTAGCTTCATTTCTG	128	(TG)18
C06405	GAAATGAAGTTATGAAGTTTTG	AGGGATTAGTGAGTTGTTTACC	143	(CA)11
C06406	ACCAAATGTCAATCAATAGATGAA	CTAGACCCATCCATGTTGTTG	131	(CA)16
C06504	CCTGAATAGAGCCTGCTTCTCC	TGTTTATTGCCCATTTGGAAA	214	(CT)6(GT)7AT(GT)2(CT)2CATG(AnT)3
C06508	CCATGAATGTTGAGTGTCTCATA	GAGCATGCTTCTCCCTCTG	186	(CA)8(GA)13
C06511	ATAGTGAAATTGCCCTAGTGGT	TATCATACTGCCCATTATGTG	114	(CA)11
C06513	TGTTGCTCTCTCTGCCTAAT	CTTTCAATCTGTTGGTGTCTAT	161	(CT)9(CA)10
C06602	ATCCTTAGATGTAGACCCTTAG	TGTCATCCAGGCAATAGAACT	137	(GT)11
C06605	TCCTCCTTAGGGACTGTACCC	GCATCACAGACGTGTCAGGAAC	131	(GT)19
C06610	CTCAGAATCAGCAGCAGGTGCC	GTTGCTAAGTTACAGACATCACCA	206	(CA)10
C06905	CAGAAACTGAGATGTGTCAAAAGTCC	ATGCCATGTTCTGATGCTCTTG	166	(GT)14
C07002	TTCTGGATGAACATACCTTTG	TGGTCAGGGGTAGAAGAGTG	81	(GT)12
C07003	GAGCCTGCTTCTGCCTCTCC	GTATTAATGGATGGATTGCA	156	(GT)25
C07004	AGTTTGAACATCCTTAAATTGAT	AATGCAGAATCCAAGAAATAGAG	118	(GT)12
C07010	CTAGTTCCATCCACATCATTG	ACAGTCCAAGTGTCCATCAAC	138	(CA)15
C07011	TTCTCCCTCTGCCTGTGTCT	GTATCTTTTATACCTTGGACCTAT	215	(CT)6(GT)15(AnTn)8
C07013	GAAGGAAGCCACCAGTAAAGT	TTCTTAGAAAGACCCGAGTA	138	(GT)11
C07102	AGTCACAGAGGGCAGTGTGG	ACATCCGCTTTAATTTGTTTC	118	(GT)17
C07104	GTAATCTCCATTCAACACAAGTGA	CGGATATAAAGGTGGGGTATT	187	(CA)9
C07108	TGCATACAGTATCAATTTGTGA	GGATAGAGTCCCACATCGG	168	(GT)10(GA)9
C07212	ACTATATTGACAAGTATGCACAAGA	GAGCCTGCCTTTCCCTCTG	183	(CA)20
C07301	GATAGATGAATGGATAAAGAAA	TTAGCATAACACTCTCAAGTT	135	(GT)11
C07302	ATCACTAAACCACCACCAGAG	AGGTAAAAGGCGAAAAGAACTT	129	(GT)9

Table 2A (cont.)

C07304	CAGTTACATATACCATTAGCCA	TGCCTCCTCTTTGTCTCCA	109	(CA)7TACG(CA)10
C07308	ACATTGGGCCTAATTTAATAGAT	GTCCTGGAGAGCTTATAGTAGACA	127	(CA)11TA(CA)3
C07403	TGCCATCTTCTGATGCTCTTG	TCGTGGTTCTTCTGGAATCTG	134	(CA)14(T3A)10
C07407	TCATTCATCAAGTCCTCAGTTAT	CTTATGGGCTGGAGGTGTGTA	121	(CA)15
C07413	TTTCACAGCAGGGAAACTGTTATG	ACCCCATCAATCAAGAGAAGTTA	120	(GT)18
C07415	AACTGTGTACTTCCTGTTCAT	ATTTAATCGACTGAATGTTCTC	101	(GT)8
C07502	CATCACCCTCAGACTGTTAGTGTT	GCATTCTTTCTGGTGGGAGGA	180	(GT)11
C07902	GACTGATGGTGGAGGGTGAG	TGTGACCAGTCTGTAAACTAC	91	(GT)14
C08103	CTTGGAAATGTAATGTGTGTA	CAGTTGTGATATTTTGTTTTCAG	91	(CA)12
C08202	ATGTTCTTAGCCAGTCATAAATC	TTTGAGGTTGGGATGTTCTCTA	203	(GT)13
C08204	TCATCTACTTCCTGTGTAGCC	GGACATAAGAGGATGTGAGAA	113	(CA)21
C08411	AAGCAGATGCTCAACCACTGT	GAGGATCGAGTCCCAGGTCAG	174	(CA)13
C08413	ACTTAACTAGAGAGCGTGTGACT	ACCTACTTGCGTGTTTTAAGG	135	(GT)13
C08601	ATATACTTTCACTCTCCATGCAA	AGAAGAGGAGTCTTTGGATG	139	(GT)18
C08608	CACAGAATACTGGAACTCATTTAG	AGAATCTTATTGGTTCGGTTTGG	155	(GT)18
C08903	AACTGACATCAACAGTCTGATAC	CGACTCTAAGATCGAGACCTC	186	(CA)16
C09004	CTACATGGAGCCTGCTTCTC	TGAAGAGGAATGGAATGACTC	138	(GT)11
C09107	CCTGCATGGAGCCTGCTTCTC	ACAAATAGGTGGTCACTTACTGAA	150	(CT)14(GT)7
C09109	TGGAGCAAGCACTTTCTATAAAC	GAGCCTGCTTCTCCCTCTG	148	(GT)16.....(GA)8
C09205	CCTCAAATAATGGAAGTGGCT	CAATCCAGTTATGAAATGTTCAC	123	(GT)14
C09210	GGTGGCTCAGTGGTTTAGCA	GGTGGTTATGATTGTACTTTCTG	149	(CA)18
C09211	TCACCTACTGAGATACTTCCAT	CTGCCTATGTGTCTGCCTTC	204	(CA)7
C09213	TTTCACCTCTGATTATATCTAGG	TGCATGGAAGCCTGCTTCTC	140	(AC)18
C09215	CCAGGAATAGACAATGCCCA	AACCCTAAGACCTTTGTAATC	255	(CA)12
C09217	CTCTGCATAATGCCTGCT	AAGACTATTTATTTATTCATAGAC	80	(TG)11
C09220	CCTACTGTTTTCTGTATTGGCA	CTGCATAAAGCCTGCTTCTCC	165	(CA)4TA(CA)8
C09303	TCTGTCAATGGATAAGTGGAT	TCCAGGTTTATTCAAGTAGTTAC	129	(CA)13
C09304	CTAGATTCATCCACGTCACTG	CCATCAACTGATAGGGAAGAT	129	(GT)12
C09305	TTGCCATCACTGATACAAGT	TTATTTCTCTTGCATAAATAGCT	181	(CA)9
C09307	TTACCCTTGGCTATCTATCTAT	CTGTTCCATCTTTTCCACCTTA	164	(GT)5G(GT)12
C09309	TGGAGCCAGTTTCTCCCTCTG	TGTTTCTTGATTTGGGTGGTA	141	(GT)15
C09310	TAGAGGATCAGGTCCCACGTC	GCAGTGCCACGAATGAGTCA	264	(CT)11......(GT)17
C09312	AACTGGAAAAATGGATAATCAG	TTGGAAAGATATTCACATTCAT	144	(CA)9
C09314	GTCACTAAATTCACGTTATTGA	CTTTTCTCAGTGTGTCTCAGAA	228	(CA)8G(CA)6
C09403	AGATTTGAACCAGGAAATTAGGAA	CTTGAGACTCTCTCTCCTCTGTCC	182	(CA)9
C09407	TGTTAATCTTCCTAATCTTCCAG	TCCACTGTTATTGGCATCACAT	104	(CA)16
C09413	TGGAGCCTGCTTCTCCCTCTG	GATCCACATCCCTGAGCTGA	202	(GT)9
C09601	TGGAGCCTGCTTCTCCCTCT	TGCTTCAAAGGACACATCAAGGT	138	(GT)17
C09607	GCTGGTTCTTTCTCTATTTATAC	TTCAAAGCTAGTCACTATTAGCA	131	(CA)13
C09609	ACTGCTGGTTCTTTCTCTATTT	GGTAAATACTTGAGGAATTAACATT	102	(CA)12
C09610	CTAGCTTGCTCCACTGAGTTCC	CAGATGCCTCCCTAAAGATGTG	163	(GT)9
C09703	GCTTCAGGAATCTAGGGACAA	TGTATTTCCTATGCAATATACC	152	(CA)16
C09805	GTGCCTGCTTCTCCCTGTCTC	CACAGCAAGTGAGAGTGAGCA	156	(GT)10
C09806	GTAGTCTGCTTCTCCCTCTCC	TTCTCATATGTGGTAACTGAGTA	208	(CA)16
C09807	GCCAAATTAACCTATATTTAGAAC	AAGGCCTCAGACATGAACTATAAT	176	(GT)6AT(GT)3
C09903	TCCACATCCTCTTATCTGTTG	AACTCAGTGGGACCTTCAATA	148	(GT)5AT(GT)11
C09912	AAGATGATAGCTTGGTCAAAGAG	GAACCAGGTAATTCTTCTATTGAA	135	(CA)8AA(CA)10
C10103	GTTGGGCTCCCTACTCAGTG	GAGTGTGGAGACTGCTTAATA	289	(CA)11
C10104	GGCAGATTTCTCAATACAGATTA	TGCTCTCATAATAGACGAATCACC	119	(CA)12
D00101	ACTCTTCTCCATCTCCCTCTGC	TCGTTGGGGTTAAAGCTCTGACC	150	(CA)9
D00103	GTACTTCCTCAGCTTTCCAATG	CTCCCTCTGCCTTGTCTCTG	177	(AT3)4....(GA)4(CA)12
D00109	TGTATGCTCAAGGATTATCTGG	TCTCTGTGCCTGTGTCTCTGGC	127	(CA)17
D00401	TGCCCTCACCAGGTGTATAGA	GTGTGAATATGATGTGTCTAGAAA	90	(CA)22
D00701	CCTGCATGGAGCCTTCTTTC	TGTATGCTCATTAACCATAGTCTT	150	(GT)17
D00704	ATGGGGGAAAGCTGAAGGAGATCC	TGTCAGACTGATAATAATGC	459	(CA)25
D01004	TCCCTGCATGGAGCCTGCTT	GAACCCAGATTCCAGTTGCTA	246	(TC)12+(GT)12
D01204	TATCCTACCTCTACACTCCTCCTG	TGAGAGTTAAGGGGGTTAATGG	589	(GT)20
D01205	AGCATGATGCCCTTCAAGGTC	GGATCTTTACCCGCATGTTCC	201	(GT)2A2(GT)16
D01208	ACTCTGACAAGGTTCTGGCG	GAGTTTATTTTGGTGGTGTC	130	(CA)12
D01210	GCCACAACTACACAAATAACTAA	TTCTACAGTGATGAATGCGAGT	213	(CA)>10
D01211	GCTTTTGTTCCCTTTAGTGA	GTTCATAGCAGCAATGTCCAC	127	(CA)23
D01212	CATAATAATTCCCACCACTACT	GGAGCCTGCTTCTCTCTCTG	133	(CA)17
D01214	ATCATTGTAAAGCAACCTCTC	TTCTCCCTCTCCCTCTGCCT	254	(CA)5(GA)6

Table 2A (cont.)

D01215	CCTGCATGGAGCCTGCTTCTC	ACGAGAGACTCCTAACTCTGGAA	260	(TG)17
D01504	CTGCTTGTAGTCTAAGTAGGTC	CTGACTGGGCACAGTGATCTA	237	(TG)5(CA)(TG)(TA)(TG)9
D01505	CCAAGGGGTATGTTGTCTATTACT	CAGCATGAAGGATCTCTGACTA	157	(GC)9(AC)13
D01702	CTCCCTCTGCCTATGTCTCTGC	TCCAACCAGAATATCAGTTCCC	450	(CT)16(GT)19
D01707	CTGATACTCAGTTCCACTCGCC	CTGGTGACAGAGGCTCAGATCC	396	(AC)10AG(AC)5
D01708	GTAGAAAGCACTGAAGACATG	ATTTGGTCACAAGATAGAGGC	279	(GT)12
D01715	TTACTGAAGTGATACTGTACCCTGC	TAACTTTCTCTTGGATGTGAAGG	192	(GC)9(AC)5AT(AC)7
D01901	TTGGGTGATAATATCTATTGCT	CCTGCTTCTCCCTCTGCCTGT	190	(CA)13
D01902	CCTACTAAAATACAGAAACG	AACTGTTAGAACTTAGACATGC	129	(GT)18
D02001	GTTCTCATAGAAGGAAGTAGGAGC	ATATTCTCTTAGGTTAGACAGCAGG	271	(AC)20
D02004	CTTCTCCATCATCATTTTAC	GTAGATATTGAAGAATGAAACA	184	(CA)17
D02005	TCTAAATATGTATATGTATGCGT	CACTTTATAACAACATATTCAAAT	119	(CA)13
D02009	TAAAGTTTCCTCATTTTCAGT	ATCCTTCTGCTTTTTGCCTAATA	143	(GT)15(GA)15
D02012	CTGAGATGTGTCAAAAGTCCTTTCG	TTGCCTACAAGATCCCTACATGCC	171	(GT)15
D02202	TTAAGCAGAAGCTCCGCTGC	AATTTTGGTGCCCACTATGGAAGCC	91	(CA)12
D02209	GCTCACCACATGATCTTTGTATTCC	TTCTCCTCTGCCTGTATCTCTGCC	180	(AC)10
D02210	GGGTCTGAATTTTGTTCAC	ACATCAGGCTCCCTTCATGG	160	(AC)11(AT)2(AC)5(AG)3
D02211	CCAGCATTACCCTGATACCA	GAATAAATCCTCCTGATTGTG	201	(CA)18
D02212	AGCCTGCTTCTCCCTCTG	CCTTAGTATCCCAGTATCAC	213	(GT)12
D02214	AAGATTCTGTGAGACAGGATCAGCG	ACTGGAGGGAAAGATAGCCAATGCC	191	(TG)16
D02919	GGTGCAGTTACTTAAAGACAG	ATGTGTTGAACACATAGTAGG	123	A15T2A10
D03202	CTGTCAGGTCACTGAGATTTAGA	CCAGGACTATACCCTCCACAT	156	(GT)15GT(GT)3
D03209	ACTGGAGTGAAAGGTTCAGGA	CTGCATGGAGCCTGCTTCT	300	(CA)3G(GT)21
D03301	CCACCACACTCCAGGTTCCA	CACTGTAAAGTAGTTGAACTTAC	231	(CA)17
D03505	GGCTCCTCCTTGGCAGAGA	CTGGACTTTGCATTCACTTTTCAG	133	(TC)4(AC)2(TC)3
D03601	GGAATCTGCTTCTCCCTCT	ACATGTGAGATGCTCAATC	185	(GT)20A(TG)10
D03707	AGAGCCTAGATGCCCATCAA	TTCACTTAGCGTAATATCCTCT	156	(GT)19
D03708	TTGAAAGAGATAAGGAGTCTGGAG	TGCAGGTCCGACTCTAGAGGAT	82	(GT)3A(GT)5
D03709	ACATTTCTGAGTGGCATGGCT	ACTCCCAAATCTTCACAAAGGAA	86	(GT)9
D03805	GTCAACAGCTTAGAAGTCACCA	ACTATTATGCTGTATGGGTGCAA	90	(AC)12AAT(AC)5A(AC)2
D03815	CTAAGATCAAATCCCACGTC	GATTCGATCTGAGTTAGCAC	172	(TG)5(TG)8
D03821	CCACCCAGGCATCCCAAGA	ATCTCAGAGAGTTGGAATCAATC	190	(AC)19
D03823	ATCTGGCTCCCTGCATGAAG	ACTTGTTTTCCCTCATATCTGTT	151	(CT)10(TG)5___(TA3)(TA4)(TA3)9
D03908	TACACCTGACACTTGTATCC	GTGCTTGTTAGTCCATGACC	94	(AC)13
D04101	CTGCATGGAGCCTGCTTCTC	GAATATGATGTACCAGGTGTGG	171	(TG)16
D04402	CCCAGGCACCCCCTTTTCTC	ATCAAGTCCCATGTCAGGCT	179	(CA)18
D04403	CTATTGATTTTTCCAAAGC	GTCTTTCATGTTTTCATATACTC	130	((GT)15
D04501	ACTAGAAGACACCAAAATGA	AGGAATCTGCTTGGATCTCT	176	(AG)4___(GT)3
D04503	GAACCTGTTTCTCCCTCTGCCT	GTCTCTCCCTTTGCCTCGTAG	158	(TG)17
D04504	GCAATCTATTAGTGGGGTCAT	CTGACTCACAGCCTGAAATGTAT	224	(TG)14(GA)3GC(GA)6
D04513	TTGTCATTGAGGAGAGTCAT	CCACTCCAGAATGTATCTAAAC	96	(CA)5TA(CA)5
D04517	TTGACTAAGGGACTCTCAG	TGGGTGGCTCAGCAGTTTA	254	(GA)3(CA)10(GA)14
D04606	CTGCTTCTGTCTCTGCCTAAT	TCCCTCTGCCTGTGTTCTCTG	280	(CT)10___(CA)13
D04609	AGCTATCTCTTCATTTGATCTATCC	CTAGAAGGACAAGTGTGTCTACTGC	225	(TG)10AG(TG)5
D04610	ATCCAAAGACAATTCAAAGG	TTGGGTCTATTTCTGGGTTCT	133	(GT)10
D04613	ATCTCACTCAGAGCGAAAGCT	CGAGTTCCAAATCTTACAGG	293	(GT)10(AT)7(AC)6
D04614	ATCAAGTCCCACATCGGGCT	GTGGTTCTTATCCTTTCTCTTATC	154	(CT)12(GT)12
D04616	TCTCATTCTTGTTTATGGCTGT	ATGCACCCTTATGTTTATTGCAG	167	(GT)17
D04617	AGGATGAGGTAGGAGTCAGAA	GCTATGCTTTGGGATGACGTG	271	(GT)14
D04702	GTCTTCCAAGTGGTAAGAGCCTACC	ATCCTCCTCTACCCTCAGAGCC	112	((CA)12
D04710	TCCCTGCATGGAGCCTTCTT	CATTCATTCTAACTTGAGTGTC	526	(GT)17
D04810	CTCCTCTCCCTCTGCCTGT	ATGAACTCTGCACTTGGCGT	231	(TG)14
D04811	TCAAGTCCACATCAGGCTTC	ACGTGGTGGTATCAAGTCTCT	189	(CA)19
D04812	TCCCTGCATGGAACCTGCTTC	ACTCGGTTTAGTTGGACTCCTTA	190	(TG)11___(A3T)12
D04813	TGGAGTCAGTAAAGCAGGCTA	TGAGTGACTGTGTTCTATCTTGT	122	(TG)10TAGTC(TG)4TCTA(TG)7
D04907	TCGATTGAGCCTCCCAAATAACT	CCATCACCCGGAGTCTGTAAT	216	(CA)13
D04911	TGATAGACACTTGGGTTGCTTCCA	ACTCTTGGGCATTTACTCCAAGGA	164	(AT)5(GT)9(AT)5C(AT)7

Table 2A (cont.)

D05005	ACATCGGGCTCCCTGCAT	ACCGTGCATGTCGCCACA	232	(AC)13
D05008	TCCCTATATGGAGCCTGCTTCT	GAAGCTCCTATTTGCCTTTCACCA	200	(CA)13
D05012	GAAACTTCATAGGCAGACAAATG	AAAGTACCTATGGTTGGAGCATA	136	(CA)17
D05101	AGGCATCAGGAAATATTGTGGGA	AGAAAACACACCCAGAGACAGG	165	(GT)16(GA)21
D05120	ACTCTGCTGTATAGACATCTTGT	AGCAGAGGACTATGGGAAATAAC	108	(TG)12
DX-4	ACATCAGGCTCCCTACATGG	CTCACTCAGGTTACTTGGCTGC	170	(CT)6(GT)7
E00402	TCACCGTTTTACCCAGTATTCC	TCCATTCCTGATCGAGTTCTG	212	(ATTT)5(AC)3(AG)11
E00409	TGCTTTTGGATGGAGCTGAAG	TGAGAGGATCAGTTTCTGTTG	211	(CT)8(GT)8
E03906	AGCAGTGAAGTTTAATGAAATAC	GCTCAGGAATTACCAGAGGAG	85	(CT)12
E03909	AGCACTTACAGGGTGTGGTCGTA	GACTTCCCAGTTGACTAAATAAGCTA	214	(CT)9
E03912	TGTGGAGTCAGCTTCAGATTC	GCTAAACCACTGCACCACTGG	150	(TC)18
E03913	AAACAAGTGGGGAGGGGAGG	CTTGATCGAGCCCTGCATTGG	117	(AG)4C(GT)7
E03914	TCAGTCCCACATGCAGCTTCTG	GTGAGACCAAATTGTTATTGTAA	202	(CT)16(GT)8
E03917	AGGGAGAACAGATACTGACTCAA	TAATCAGCCTCTAAGGATTCTGG	216	(AG)14
E03920	CTGTGTGAAGCCTGCTTCTC	AGCCAGTCATGTGCCCTTA	132	(CT)9G(TC)3
E03922	CACATTTTACATAAAAATAATATGCCA	CAGTGCATGGAGCCTGCTTCTC	192	(AG)17
E03923	CTGCATGGAGCCTGCTTCTTCC	GTTTCAGCATCTGCACCAGGAT	172	(CT)14G(TC)3
E04001	TCAGCATGGAATCTACTTGAG	GAATGTGAGTACAAAGGTAGG	76	(CT)11
E04007	GCTCATTGTGATTCCTTAAAACAG	CTGGGGTCCGGGATGGAGT	202	(GA)5A(AG)15
E04008	GGTAGCCTGCTTCTCCCTCTG	ACCAGTGATTCCCTTCACCTG	143	(CT)12(GT)5
E04019	GCCCTCACTGGACATCTTTATT	TGGAGCCTGCTTCTCCCTCTG	116	(GA)13
E04021	CAGTTTGGAGTCTGCTTCTCCCT	ATCACCTGAATTGCAGTTGTCA	182	(CT)10
E04104	ACTAGGCATCTCACATACATTATT	CCTGCTTCTCCCTCTGCCTAT	109	(AG)12
E04105	CCTGGAATGGAGCACCATGTC	ATACTTATGTCCCTGGCTCTG	168	(CT)8C2T2(CT)6
E04107	CTCCCTCTGCCTATGTCTCTG	CCAAGCAGTTTTACCACGATA	110	(CT)12
E04108	CTTCTCCCTCTGCCACTTC	TTCTTTATTTGACAGGGAAA	98	(CT)10......(CT)6
E04401	CCTGGCATGGAGCCTGCTT	GTTTTTAGGTCTACACTTCTGAGT	122	(CT)9 (GT)3
E04402	TGAATCATTATGGTCCTATCGTTC	TAAAATGCAAGTCTTACCAGAGGAA	111	(TC)13
E04403	TGCATGGAGCCTGCTTCTC	CCTTTCATTGAATATCTGTCAT	123	(CT)11
E04404	GCCACATAGACACTTGGTGTT	CGGGATGGAGCCTGCTTCTC	114	(GA)12
E04407	GGAGCCTGCTTCTTCCTCTG	CACTAGTAGCTTTATAATTGTGCT	124	(CT)14G(TC)4
E04408	TGCTTCTGGAAACTGCACAT	TGCATGGAGCCTGCTTCTC	144	(AG)12
E04409	AGCCTGCTTCTCCCTCCTC	GTTTTTAGTCTACACTTCTGAGTAA	111	(CT)9(TG)3
E04411	GAGATCGAATCCCACATCAG	CCTACTCTTCCACCATTTTGCC	166	(CT)11
G00203	CTCTGCCTATGTCTCTGCCT	TGTATGTCTATTTTTGTGCCAGTA	164	(TC)13
G00402	GTTTGAACCCCTGCCATAGGTA	CGGAATCGAGTCCCACGTCA	175	(CA)5(GA)20
G00410	TGGAGCCTGCTTCTCCCTCTG	GCCAACTCTTTACATCTGTGCTA	148	(CT)11
G00501	ATGCCCACGTCAGGTTCTCTG	GTTGTTCCAGTATTCATTCATTC	171	(CT)11
G00504	CCTGCTCAGCAGAGAGTCTG	GATTGGATTATTTGTTCTTGG	161	(CT)14
G00508	AGTGCGTGGAGCCTGCTTCT	GATGTACTGGCCCATCATTCT	196	(CT)14
G00512	CAGGGCTCAATGAGTGATGTTA	TCAAGTCTTGCATCGCACACC	158	(CA)15
G00602	CGAGCTGCTCAACCGCTCAAC	TGGAGCCTGCTTCTCCCTCTG	187	(GA)19
G00605	TTCTCCCTTTGCCTGTGTCT	GTCTATGAGAGCACCAGGTTCA	190	(CT)11
G00703	CTTCTCCCTCTGCCTGTGTCT	AAGTTGTGTATTGATTTCATTCTG	206	(TC)6T3(TC)7CA(CT)3G(TC)3
G00704	GGTCCTCTGAATCCCTGTCTAT	GTGGAGCCTGCTTCTCTCTTTG	225	(CT)9T(TC)5.....A17
G00707	CTTCTCCCTCTGCCTATGTCTCTG	GAAGGCTTAGCAAGAGTTGAAGA	189	(CT)13GACTATCA(TA3)2(TA5)(T2A10)(TA6)2
G00708	CCTCTCCCTCTGCCTGTGTCT	ACCTCTGAATCAGGAAATGTAACT	132	(TC)12
G00709	ATCGAGTCCGACATAGGGTTCACT	AAACAGTGTAAACAACATGCTACC	152	(CT)12(GT)4(CT)3
G00712	ATCGAGTCCCATGTTGGGCTCC	TGAGCAGGGGCAATAGGAGACTTC	226	(CT)9G(TC)3ATG(A2T2)(A3T)2(A4T)(C2T2A8)
G00713	CTGGATGGAGCCTGCTTCTC	GCGTATCTAGTGATGCCACTTCT	194	(CT)10T(TC)3ATG(A2T)(A3T)2(A4T)(CT4A4)
G00801	TGCTTATGCGTACTCTCTCTCAA	TCCCTGCATGGAGCCTGCTTC	184	(CT)12GAG(TC)3ATG(A2T)(A3T)2(A4T)(CT3A9)
G00810	CTCCCTCTGCCTACGTCTCTG	AGAAGTTACTGTGTCCAAGTACAA	152	(CT)17(GT)4ACTATCAT(A3T)2(A4T3)(A11T)

Table 2A (cont.)

G00812	CTGCTTCTCCCTCTGCCTGTATC	AGGAACTGGCATTCTACATTAGCA	198	(CT)11(GT)3CTCATG(A2T)(A3T)2(A4T)C(T3A6)
G00903	TGCTTCTCCCTCCTTCTGTGT	ATTGTGAAAATCCCTCCTTAGAAAT	142	(CT)11
G00908	CCTGGTGCATGGAGCCTGCTT	GATGCTGCAATGAACACGAGAGCT	145	(CT)12
G01006	GAGCCTGCTTCTCCCTCTG	TTTATTCTCCCTGTGTTCTT	113	(TC)18TATCA(A3T)2(A5T2)A10
G01109	TCCTTCTGCCCCTCACCC	AGCCCAAGTTATAGACAATGAT	112	(CT)18C2T2(A5G)2A9GA4
G01204	CATAGGGCTCCCTGCATGG	AGCCATTTGTATGTCTTCTTTGTA	226	(TC)17(TG)3TCTC3ATG(A2T)(A3T)2(A4T)
G01303	CTGCTTCTCCCTCTGCTTTGT	GTGCTAGATGGGGGCTTCCTC	118	(TC)17GTG(A2T)(A3T)(A4T)CT3A8
G01305	TAGCTGAATGAAAGGGCTGATAG	TGCTTCTCCCTCTGCCTGTGTC	181	(TC)16___(TA3)2(T2A6)(TA2)(T2A10)
G01406	ATCAAGTCCCACGTCAGGCTTCC	ATTTCGAGTGTTTCTTCGAGAAGTT	161	(CT)15___(A2T)(A3T)3
G01506	TGGAGAACCAAATTGAGTCCT	GAAATCCACATTATATGAGGTTAAAC	155	(TC)16(GT)2
G01509	CTAATGTAACATTGTGTGACAACTACA	CATGGAGCCTGCTACTCCCTCT	110	(GA)9 G2 (CA)(GA)2
G01511	CCTTGCTCACCATATCACACA	TTCTTTCTCTGCCTCTGTCCT	153	(GA)5 GA3 (GA)5
G01515	TGCTTCTCCCTCTGCCTATGTCTT	GTGCAGGGCTCAATGAGTGATGTT	133	(CT)16
G01617	TGGGATGGAGCCACAAGTCA	CTTACGACTGTTTTCTCAACCTG	240	(CT)10
G01621	CCACTCCCATCTCTGCTCAT	CCAACGACTGAAGCTGTCAT	134	(CT)4CA(CT)6
G01705	TGGAGCCTGCTTCTCCCTCTG	GGGGTTGCCTCTTCCTCCTTT	125	(CT)9
G01707	TCATTGCCAGACCAGGTGTC	GTGCATGGAGCCCGCTTCTC	159	(GA)9
G01709	AGGGAAGACCCGTGACCAT	GCTTCTCCCTCTGCCTGTGTC	258	(GA)10
G01713	ACTAGAACTACAGATCAGTCC	GAGAACAATGGCAGTTGTCT	187	(CT)8
G01715	ATGGAGCCTGCTTCTCCCT	GGGGTTGCCTCTTCCTCCT	128	(CT(9
G01717	TGGAGCCTGCTTCTCCCTCT	CTGCATTTCCCTGATGACAT	172	(CT)11
G01804	CCAAGGATCAAGAACCACGTC	GATGCACTCTCCAGTTGAACTA	168	(CT)14
G01807	AGGATCGAGTCCCACATTGG	TCAGTTAGAGCATGAATCTTGTC	205	(CT)2GC(CT)12TT(CT)4(GT)4
G01811	TATGAGTTGGGCTCCTGGTC	CTGGGACAGTAACACACATTAGT	197	(CT)16TT(CT)3
G01817	AGTCCTGTGTCAGGCTCCAG	ATAGTGCATTCTTTTCAAGGAC	152	(TA)6
G01901	CTCCCTGCATGGAGCCTACTT	CTAGAGTTCTCTCAAATCTGTCA	130	(TC)11 (GT)2 (TC)2
G01903	AATTAGCAGGGAGTCTGTTTC	GGTACTTGGGTTTTAGAATAT	165	(CT)4T2(CT)5
G01905	TGAACCCTGCTTCTCCCACTG	ACGACTTGAGCCACCCAGGTA	169	(CT)9 (GT)2 (CT)2
G01906	GAGTCTGCTTCTGCCTCTG	CTGTACACTCTAAATGGGGTCATT	152	(CT)9(A3T)2CT2A6
G01918	TGTCTCATTCTAGCTGCTACATT	CTTCTCCCTCTGCCTGTGTC	106	(GA)18
G01920	TGGAACATATCTTTTTGGGTGACC	CCTGCTTCTCTCTCTCTGCCTGTG	233	(CT3)23__(CA)6(GA)7
G02002	AGGATCATTGGCTAGACAAAC	TACATAGTTGGGATCGAGTCC	248	(GA)10
G02007	TCCCTGCATAGGGCCTGCTT	GAATAAAACCTAGACTGGCTGAAG	128	(CT)2GC(CT)7
G02106	CATGGAGCCTGCTTCTCCCTCT	AAGGCAGATGCTCAACCACTGA	159	(CT)9
G02107	CTGCCCAGAGAGAGTCCTCCAT	TCGAATCCCATGTCGGGCTC	189	(GA)10
G02108	CATGGAGCCTGCTTCTCCCTCTG	AGAATATCTTGGCTGCAATGCTT	146	(CT)13
G02111	ATTGGAACTATGCAGGCTAT	CCTGTGTCTCTACCTCTCTGT	163	(CT)13
G02202	GGATCGAGTCCTGCATCGAG	CTGAGCCAAAGGCACTCAACAG	177	A15(GA)9
G02204	ATCAGGCTCATCCCGCATCAG	ACATAAGGAACTTCTCCATCCAT	200	(CT)9
G02501	GAGCCTGCTTCTGCTTCTGCC	GCCTATGGTCTTATGGGTGTTCC	132	(CT)9
G02504	TAGAGGATCGGGTCCGGCCTC	TTCACATGGTCTTCCTTTTGGGT	197	(CT)16
G02506	GCAGAAACATACACTCAGTAGG	CCCTCTGCCTGTGTCTCTACC	179	(GA)14
G02509	GAGGATCAAGTCCCATATTG	GTAGGCAGCGTACAGATGGAT	135	(TC)9
G02512	CGCTCATGCAAGTCATCACAT	ACACTCTGGTGCAAGCGACTC	125	(CT)15
G02513	CATTTCTCAGCATGTATTATAGAT	GTCGGGCTCCCTGCATAGG	120	(GA)14
G02602	TACTCTGGATGCACTCATAAGG	TGGCTTAAAACCTACTCCTCAG	123	(CT)14
G02610	CTTTGCCAGTTATGGGTCTGTG	TGCCTGTGTCTATGTCTGCCA	132	(GA)16
G02616	GCCTACTTCTCCCTCTGCCTATG	CTGCTTCTCCCTCTGCCTTTTC	163	(CA)2(GA)10
G02619	CCTGCTTCTCCTTCTGCCTGT	TTAGTTTTCACCAACTGTAGGG	154	(CT)9
G02620	CTGCATGGAGCCTGCTTCTCT	GAATTGTAAGTTTTCAACTGCC	144	(CT(9C(CT)4
G02702	ATCACAACCTAACCAAAAGGCT	CTCTCCCTCTGTCTGCCACTCC	142	(GA)12

Table 2A (cont.)

G02704	ACCCAGGTGTCCTTCAAAATGT	GCTCTCCCTCTGCCTGTGTCT	206	(GA)9
G02709	ATGGAGCCTGCTTCTCCCTCT	TCAGCTATAAATTCAACTGGCTTA	151	(CT)14(GT)(CT)2
G02710	GGCACGTTAGTCTAGTTCTCTG	TAATCAGGTTCTTGGAGATGAC	139	(GA)8AT(GA)
G02712	CCAAATTCAGGATTTCTGACTCC	ATGGAGCCTGCTTCTCCCTCT	161	(GA)12
G02806	GCAGCCAATATGACATCATCC	TACATGGAGCCTGCTTCTCCC	161	(GA)8
G02807	TGCATGGAGCCTGCTTCTCCC	GAACAGCTTTTGCAGCACCC	175	(CT)11
G02812	TAGCTGTGAGCTGGGTGTGGA	GGCACTTCACTTAATCTTTGAGT	114	(CT)7
G02813	CGAGGATCGAATCCCACGTC	TCATTTGTCACTTATTAGTCCAC	174	(CT)2GC(CT)9(GT)3
G02814	TGCTGCTTTATAGTAAAAATG	CCTGCTTCTCCCTCTGCCTAT	265	(CA)5(GA)12
G02815	TCCTGCTGAATATGACGTTCA	AAGGGAGGGGAAACGACACAT	154	(CT)15
G02817	ATCGAATCCCACATCAGGCTC	CACAAATGTAAACTGGGTATATT	177	(CT)9
G02819	ACACTCAGCATAGAGTCTGCTTG	CACCAGGTTGGAAATGAATAAG	154	(CT)13
G02821	CCTGCACAGAGCCTGCTTCTC	AAACCACTGAGCCACCCGGACT	153	(CT)14
G02902	GATTGAGTCCCACATCAGGCT	AGCTGTGTTTATGACTACACATG	241	(CT2TG(CT)6.....(CT)6
G02903	TAGAGCCTGCTTCTCCCTCTG	CCAATTTGAAGGATTCATCATT	146	(CT)2GT(CT)7GTAT(CT(8
G03001	TCCATCTGCCTATCACACCACT	TGAGCACTGGATGTTATATGCAA	199	(TC)9
G03006	ATCTAATCCCACATTGGGCTC	ATGGGGAGTCATCAGACCAGG	171	(TC)13
G03011	TAGGCTTCCTGCTCAAGACAG	GGATGGAGGAGAGGCTTGTTA	209	(GAT)6.....(TC)9
G03012	CTGCTCTCTTTTCGCTCACTC	TTCTCCCTCTGCCTGTGTCT	141	(GA)17
G03013	ACTGAGATGGGAAGGGGCAGA	CTACATCGGGCTCTATGCTC	83	(GA)8
G03016	GAGCCTGCTTCTCCCTCTGC	AGTCCTGTGATTAGTTCTCAGAC	106	(CT)10
G03017	TCCTCCCAACATTCTACAATGAA		134	(GA)3CA(GA)9
G03018	TGCTTCTCCCTCTGCCTGTGT	CCTTCTGGATCTGCTTTTACTAT	203	(CT)13
G03019	CCACTCAGATGTCCCTATACTAT	AAACAGGATCGAGTCCCACA	212	(GA)13
G03104	TAGCAGACAAACCCCAACTG	GAGCCTGCTTCTCCCTCTG-	167	(GA)13
G03109	CTGCATGGAGCCTGCTTCTT	TCTTATTCAAATCCTCCTGATTAT	153	(CT)9
G03111	CCTGCATGGAGACTGCTTCT	TGTTTCCTCACTTCTTACTGA	218	(CT)21
G03601	GACACCAGGTTGATTATCATT	TGGAGACCTGGGATTGAGTC	166	(GA)10
G03901	ATCACACCCTGGGCTGAAGG	TGGAGCCTGCTTCTCCCTCTG	174	(GA)14
G04801	AGGATGCCCAGTTACATTTGAA	TGATGTTTGATGTTCACGTTGAT	208	(GA)18
G05002	CACTGTGTATGTCCTCTTATTAAG	CAGGAGTCTACTTTTCCTTCTG	170	(GA)30
G05602	CACTAAACCACTGAACAACCT	GTCCCACGTCAGGCTCTCTG	158	(GA)9
G05602	CACTAAACCACTGAACCACCT	GTCCCACGTCAGGCTCTCTG	158	(GA)9
G05604	TGCATGGGGCCTGCTTCTC	CCTCTTCATACTTCAGCAAGTG	169	(GT)9
G06202	CCCTTCTCTGTCTTTGAGAGT	AGCCTGCTTCTCCCTCTGCC	144	(GA3)C(AG)9G(GA)5
G06204	CTTCTCCCTCTGACTGTGTCT	TCCCTCAAAATTCAACATACAA	168	(CT)11(GA)3(CT)2
G06208	CCTGCTTCTCCCTCCTCCTG	TCCACAAAGCTCCCTACTCAT	163	(CT)10
G06211	CACTGGGCGTGTAACCTGCT	CTGAAATGTAAGTGCAAAGGAA	172	(CT)12...(A3C)8
G06219	CTAATATCAAAAGGTTATCCAC	CATCTTCCTCTGCCAGTGTC	267	(GA)11
G06221	GGATAACCAGGATAATTTCCTAC	AGAGAGGCCCACATCAGGCT	156	(AT4)(AT3)3(GA)13
G06222	CTGCTTCTCCCTCTGCCTCT	AATTTATGGAAATGTTCCCAA	150	(CT)17T(TC)3
G06224	GAGCCTGCTTCTCCCTCTGCC	ACCCATGTATGAGCCCATTGA	137	(CT)19
G06303	CAGGTGCTGCAAGAGCTTAGA	CTTCTCCCTTTCCCTCTGCC	176	(GA)17
G06305	GTCACGTCTTCAACCCTTCCT	ATTGAGTCCCCCATCAGGCTT	215	(GA)14
G06316	AGCCTGCTTCTCCCTCCTC	CCACACCTCACACCGTGTA	125	(CT)15
G06320	ACTGGCAATGGGTCTGAAAATAG	CTCAGTTATTTGTGGGCTCTTT	216	(GA)13
G06401	TGCTTCTCCTCTGTCTGTATCTC	CAGGTCCCCCTACACTAAGTG	133	(GA)10
G06402	ATGAATAGCTTGTGCATCAGTGCATT	TGCTTCTCCCTCTGCCTGTGT	132	(GA)13
G06407	CCATCAAACTTTTACAGTGAA	GGGTCTGCTTCTCCCTCTCT	163	(GA)12
G06407	CAATCAAACTTTTACAGTGAA	GGGTCTGCTTCTCCCTCTCT	163	(GA)12
G06502	GTTAGGCTCTCTGTTCAGTGG	CGGTGATACCTTCCTCATCAT	146	(CT)9
G06601	TGTGGAAACTGCTTACAATTTTC	TGCGTTACCTTACAAAGTTATTG	158	(CT)17
G06602	TCGAATCCCACGTCGGGCT	ATGTTACAATGATCTGATTTATTCT	236	(CT)5GT(TC)12
G06603	CATTCAGATGCGGGAGTTTC	CCAGGTGAGGTCCAGTTGTG	211	(GA)9
G06607	CTTCACAAGGTTGCACAAGAG	CTGCTTCTCCCTCTGCCTGT	159	(GA)14
G06608	TGATAGGACACTTAGCAAAGGCT	GAGCCTGCTTCTCCCTCTGC	194	(CA)2(GA)12
G06619	ACAACCTACAGAATGGGAGAA	CTTCCACAGCCTTTTATTGT	196	(CT)10
G06701	GCTTTTCACCCAACGACTTAGA	AACTCTGTGGCTCAGCAAGG	211	(GA)12
G06703	CTTCTCCCTCTGCCTGTGTC	GCGTCTATAATCATCAGAAAT	159	(CT)11(GT)4
G06705	CTCTGCCTGTGTCTCTGCCTC	CTATACACATTGAGAAATGGCA	168	(TC)13

Table 2A (cont.)

G06706	ATCGAGTCCCACGTCAGGCT	TTATTTATTTATTCATAGAGATGCA	98	(CT)13...ATG(A2T)(A3T)2(A5T)
G06707	GGTGCATGGAGCCTGCTTCT	TGCCAGTTCAGTTTTCAAAGTT	147	(CT)17(GT)2
G06710	TTCCTTGTTTCTATTCTCCTC	AACCCGGGATTGAGTCCTG	167	(GA)14
G06713	GAGATCGAGTCCCATGTCAG	CTTTGAGGAGATAAATCTTTCTA	225	(TC)20
G06714	ATCAAATCCCACATCGGGCTC	ATTAGTTCAAACCTCGCCAATG	163	(TC)12
G06715	TTGATCGAGTCCTACATCGG	TCTTGGGTAAACTACTTAACTT	174	(TC)11
G06717	TGGAGCCTGCTTCTCCCTCT	CCTTATTCAGATTTACCTGTTTG	147	(TC)8(TG)3
G06801	AGGGACGTGCTTCTCCTTCTG	CAATGATTATGGTTTGTCAACTT	162	(CT)17
G06805	GACACCCAACCGCTGAGCAC	GAGCCTGCTTCTCCCTCTGCC	168	(CA)3(GA)9
G06901	GGCAGCTTTGATGACTGATTTGA	AGTCCTGTGTCAGGCTCCCT	211	(GA)17
G06908	GGAACACGTTAATTCATAAAAATGAT	ATGGAGTCCCACGTCAGGCTAC	205	(GA)3CA(GA)10
G06909	GGCAGCACTAAACCACTGAG	TGCTTTGCATCTTCCATTTT	209	(GA)15
G06910	CTGTGCTCAGCGGGGAGTCT	TTATCTTAGAGTGATGGAGAGTGG	101	(CT)15
G06914	GGAAGATGTTGTCTCTTATCA	GGGTAGGGGTTTTGTTTATGG	159	(GA)20
G07001	CTACATGGAGCCTGCTTCTCC	TCCCCACAACTTTATGTCCTC	129	(CT)11(GT)4
G07002	CCTTCTCCCTCTGCCTGTG	GCCACTGATTTATTCTCTGTA	197	(CT)11
G07004	TGCTTGCTCTCTCTCAAATAA	GTGCATGGAGCCTGCTTCT	181	(GA)10G(GA)3
G07005	CCCTCTGCCTGTGTATGTGTC	ATGGCAGCAGGGAGTAGTCCA	134	(CT)12
G07006	CAGTGGGGAATCTGCTTGAG	CATTTCACTACATATACAGGTGTCA	150	(CT)11......(CT)10
G07007	AATACCTGGGTAAACATTTA	GGGATCGAGTCCCATGTC	156	(GA)11
G07008	GTGCATGGAGCCTGCTTCTC	AATGTACCTGTCCCCTTTTG	127	(CT)13
G07301	GCATTCACCCAATAGTCCTTG	GCTGCTTCTCCCTCTGCCTAC	135	(GA)13
G07308	TCTCAATTTGAAAAGTTTATAGTC	TTCTCCCTCTCCCCCTATC	174	(TA3)9___(GA)5___(GA)7
G07310	TATGCTTCTCCCTCTTCTCTG	GGTTTCTCTCCTTGATTTGTAAG	159	(CT)14
G07312	CTTCTCCCTCTGCCTGTGTC	TGCTAAACTCAACTCTCCTAA	123	(CT)14
G07314	CCATCAGTTTGTTCTCTATCA	GAAGCCTAAGTGAGGAGTAG	224	(CT)11
G07402	GGAGCCTGCTTCTCCCTCTG	TATCGTGCCACACTGCTGAAT	244	(CT)11T2(CT)5
G07406	AATTTAGTCGAAGAATGAAAGATG	GAAATAGCCTTAAAAGCAATGTA	221	(GA)14
G07407	CCACCTGGGCTGCACTGAAGA	TGGAGCCTGCTTCTCCCTCTG	134	(GA)10
G07408	TGTCACTCTGCTCTCCACCTG	AGTGCCTAAAGTTCTTCCTATTG	135	(CT)14
G07410	ATCTCCTTCTGCACTCCTGCT	CACGTAAGGGATGAGTTCAGGT	147	(TC)8(TG)2
G07413	CTGGAACAGAACCCACAATA	ACGAGATCAGTCCCACATCAG	231	(GA)24
G07414	TCCCTGAAAGGGGCATTTAAGACC	AGCCTGCTTCTCCCTCTGCCTATG	128	(GA)11
G07420	TCAGGAGGTGAGTTGCTTGGAG	CGGTGCATGGAGCCTGCTTCT	162	(CA)3(GA)16
G07502	CTCCCTCTGCCTATGTCTCTG	ACAGCCCTGTTTACCGAGGTG	255	(CT)14
G07503	CAGGAAACTGCTGGACTTGTGCT	TGCTTCTCCCTCTGCCTGTGT	126	(GA)15
G07504	AGTTCTGGAGGCTGGGAAGTC	GGTGTGAAATGGCTCTTTAGATA	215	(CT)23
G07505	TGCATGGAGCCTGCTTCTC	AGCAGGTTACTCTTAGTGACTCC	138	(CT)11
G07506	ACTTCTCCCTCTGCCTGTGT	TTCCAGTGTATGTTGATTGAA	124	(CT)13
G07507	ATGGAGCCTGCTTCTCCCTCT	GTTTCCTGCTCTCCTACCTGG	163	(TC)9
G07508	AGCCCTGCTTTCTCCCCTCT	GATTTTGATTTACATTCACAAGTACA	98	(TC)>10
G07510	AGGCATCCCTTACTTACTTACTTG	TCCCACATCAGGCTTGCTGTAT	152	(GA)9
G07701	TATTCAAGCCATTGACGGATTG	CATGGAGCCTGCTTCTCCCTC	247	(TG)2(TC)2GCC(TC)16
G07703	CTGCTTCTCCCTCTGCCTATG	TTTCCAACATTATGCTATGAT	198	(CT)14
G07704	AGCCTGCCTCTCCCTCTCCA	AGAGTCACAATGCAACCCCACAA	246	(TC)24
G07706	GGTGACACTATACTGAACCTTCT	TCTTTCTCCCTCTCCCTCTGA	116	(GT)11
G07707	CTCCCTCTGCCTGTGTCTCTG	AATTTTTATGTGTCCTGGTTCAGCC	202	(CT)9
G07709	CATTTCGCTCATGTGCCTGACTGA	CATGGAGCCTGCTTCTCCCTCTCC	147	(GA)16
G07710	GCTTCTCCCTCTGCCTCTATCTCT	ATTGATCCCGGATTTTGGTAATA	175	(CT)9
G07711	TAGTTCTTTCTGCCCTTCTCC	CATTTCCAATCCATTTAGAGA	149	(CT)9
G07712	CTGCATGGAGCCTGCTTCTC	TCAGACGCTCAACCAACTGAG	179	(CT)9
G07713	CTTGAAGGCGGCTGTTCTTG	TTGGACTTTCTCTCCCTCTCCT	234	(GA)16
G07803	CAGCATGGAGTCTGCTTGTC	AGCTAAACATTTAACCAACTGAG	219	(CT)14
G07804	GGGTAGAACTGACATTCTTT	CTGTAGGGAGCCTGCTTCTC	133	(GA)7CA(GA)8
G08002	GGTATGGTCCTGGAGACCTG	CTAATTGAGGAGATAGGATACATAATA	153	(CT)17
G08003	GTCAGCTTAGCCATTGAAGAAT	CCTGCTTCTCCCTCTGCCTC	174	(CA)2(GA)15
G08003	GTCAGCTTAGCCATTGAAGAAT	CCTGCTTCTCCCTCTGCCTC	174	(CA)2(GA)15
G08004	GGCACAACACTCTGAATTATTAG	CACTCATTTATTGCCCTACTTTTA	175	(GA)15
G08005	GGTCTTCACTGCAAGGGAACT	CATCAGATACTCCAACATTCAG	190	(GA)20
G08007	CAGAGTATCCTTGCCTGTAG	GTGCCTGGAGCCTGCTTCT	139	(CA)3(GA)12
G09201	TGGTACTGTAGCTTTGAAGAT	TCTGTGAAAGACACCCTATTTA	173	(CT)14

Table 2A (cont.)

H03501	TTGCCTTCTGGGTGTATTGACTT	GAATGTGGTTAGTAGAATTATACAG	300	(AT3)10(AT2)2AT
H03502	GATCCTGATTGTTCTTGAG	GGCATGGAGCATACTTCA	155	(AT3)4
H06601	TGCTTCTCCCTCTGCCTGT	TGGTGAAAGATTAGCCTGTGGA	125	(AT3)5___(AT4)(AT3)2
H06602	AAGTCCCACGTCAGGCTC	ACGTCACCACAACCATCTAA	165	(AT3)12
H09205	CATTTGCTGAGTCAAGGAATTCT	AGTTACCTGGAACTTGTCAGAA	200	(AT3)12
H08505	TGCATGGAGCCTGCTTCT	CTTCTACACATGTTGTCCCT	160	(AT6)(AT4)2(AT3)13
H09208	AGTCCAGCATCACCGTTTGT	GAGGCTTATTTTCTGTCCAGTT	144	(AT3)9(AT4)
H10101	TCAGGCTCATGGGATTGAGACTTC	TGCCATTGCACAGGATATAGGTCCA	305	(AT3)11
H10103	TCCACACTCAGTGCAGAATCTGCTT	TGTGAGACCGCAGAATACAGTACTC	141	(AT3)11

Amplification reactions were carried out under standard PCR conditions described above using the annealing temperature indicated for each locus or a touchdown PCR protocol (Don, R.H. et al., *Nucleic Acids Res.* 19:4008 (1991)) was established. The variability of these loci were evaluated using the dog panel. For each locus, 5-10 dogs were studied in each breed. The number of alleles observed are presented in Tables 3A and 3B.

Table 3A

Marker Locus	Mixed Breed	Cocker Spaniel	Labrador Retriever	German Shepherd	Beagle
D00101	3	2	2	2	3
D00401	5	4	3	6	3
D01205	4	2	4	4	4
D01902	6	4	6	3	4
D02001	4	3	3	2	4
D02005	3	3	3	3	3
D02011	3	1	3	3	2
D02012	5	4	3	3	4
D02202	4	1	2	3	4
D03709	5	4	3	4	2
D03805	6	4	4	3	3
D03908	4	4	3	5	4
D04403	2	3	1	1	3
D04702	3	1	3	2	3

Table 3B

Marker Locus	Doberman Pinscher	Siberian Husky	Scottish Terrier	English Pointer	Greyhound
D00101	3	2	2	3	2
D00401	3	6	5	5	5
D01205	2	2	1	3	3
D01902	5	3	4	4	7
D02001	2	4	3	2	3
D02005	1	3	2	3	3
D02011	2	3	4	5	2
D02012	3	3	4	4	3
D02202	1	3	2	2	1
D03709	4	6	4	5	4
D03805	3	7	4	5	4
D03908	3	8	3	4	4
D04403	1	3	2	3	3
D04702	2	3	2	3	2

In general, all of the microsatellite loci tested displayed variability within and across breeds. While 9 cells out of 140 (6.4%) in Tables 3A and 3B were monomorphic, these were scattered though 6 different microsatellite loci, which were quite polymorphic in other breeds. The maximum number of alleles detectable by this analysis for a locus in a given breed was 8, in the case of locus D3908 in the Siberian Husky. The percent heterozygosity observed at each locus in each breed is presented in Tables 4A and 4B.

Table 4A

Marker Locus	Mixed Breed	Cocker Spaniel	Labrador Retriever	German Shepherd	Beagle
D00101	20	0	0	0	90
D00401	100	100	100	88	25
D01205	70	50	0	22	64
D01902	100	100	100	11	36
D02001	40	86	57	50	33
D02005	90	29	38	22	27
D02011	38	0	25	44	18
D02012	0	17	33	0	33
D02202	20	0	0	0	0
D03709	20	100	75	89	50
D03805	100	50	50	30	67
D03908	100	100	100	88	100
D04403	100	100	100	100	100
D04702	22	0	80	0	30

Table 4B

Marker Locus	Doberman Pinscher	Siberian Husky	Scottish Terrier	English Pointer	Greyhound
D00101	60	0	78	86	38
D00401	33	50	86	67	100
D01205	60	44	0	86	25
D01902	100	63	100	100	100
D02001	100	57	25	50	13
D02005	0	50	77	71	100
D02011	20	33	44	43	50
D02012	0	50	17	40	0
D02202	0	0	17	17	0
D03709	100	78	100	86	100
D03805	100	67	100	80	29
D03908	33	44	100	100	100
D04403	30	50	56	14	29
D04702	67	20	33	60	40

No heterozygotes were observed in only 21 out of 140 (15%) of the loci/breed combinations studied. At the same time, 30 out of 140 (21%) cells showed 100% heterozygosity. The mean and standard deviation of heterozygosity observed for each locus across different breeds, as well as the mean and standard deviation of heterozygosity observed within each breed across different loci are shown in Figures 1A and 1B, respectively. The breeds studied show a mean heterozygosity ranging from 36 to 60% across different microsatellite loci with considerable standard deviations. Among the loci studied D03908, D01902, D03709 and D00401 showed the highest mean heterozygosity across breeds of 87, 81, 80 and 75%, respectively. The number of repeats in the reference clone in these loci were 16,18,12 and 22. The least informative loci across breeds were D02202 and D02012 at 5 and 19% mean heterozygosity, respectively. The number of repeats in the reference clone in these loci are 12 and 15, respectively. Correlation analysis did not reveal any significant linear relationship between the number of repeats at a locus and its overall observed heterozygosity (r=0.22).

Figures 2A-2D show the results from typical gels used to evaluate the alleles in gathering the data as described above. Amplification products of DNA from various different breeds at the locus D02011 are shown. Figures 2A-2D represent different gels, run under similar conditions. Note that the molecular weight marker identified in lanes marked M is the 246 bp band of the 123 bp ladder (Gibco-BRL, Gaithersburg, MD). The size of the amplification product in the reference clone was 238. The different alleles are easily identified, with PCR products separating in sharp and well resolved bands, near and below the 246 bp marker. Some non-specific amplification products can be observed, especially in cases with higher template DNA concentrations; however, these do not interfere with correct typing.

The results indicate that microsatellite loci containing CA repeats are abundant and highly polymorphic markers for the canine genome. These findings indicate that such markers hold great potential for use as linked markers for genetic defects in pure bred dogs.

The estimate that there is one useful CA repeat every 31 kb in the canine genome is in good agreement with one every 42 kb estimated recently by others (Rothuzien, J. et al., *Theor. App. Genet.* 89:403-406 (1994)). In the above-described study, a secondary screening was carried out and only very strong hybridization signals were accepted as positive, which resulted in elimination of about 20% of the primary positives. It thus appears that the estimate of the minimal CA microsatellites

frequency in the canine genome is accurate. These estimates have practical implications particularly, since most cosmids have insert sizes in the 30-40 kb range, the likelihood of finding a useful CA repeat in a cosimd clone harboring a gene of interest is high.

SPECIFIC EXAMPLE II

Materials and Methods

Patients and pedigrees. The patients and pedigrees used were primarily those used and described earlier (Yuzbasiyan-Gurkan, V. et al., *Genomics* 15:86-90 (1993)). Briefly, pedigrees of American Kennel Club registered Bedlington terriers were associated with the help of Bedlington terrier (BT) breeders. While all of the pedigrees have a family history of CT, not all had a symptomatic proband at the time of pedigree ascertainment. Diagnosis of dogs as to whether they were affected or unaffected with CT was made in all cases by quantitative copper assay from liver biopsies performed at 1 year of age or older by criteria earlier described. DNA was extracted from peripheral blood samples collected in acid-citrate-dextrose as anticoagulant as described (Yuzbasiyan-Gurkan, V. et al., *Genomics* 15:86-90 (1993)).

Microsatellite analysis. The microsatellite markers used in this study were developed as described in Specific Example I. Standard conditions used to amplify each marker locus in polymerase chain reactions (PCR) were as follows: 25-50 ng of genomic DNA as template in 25 μl of PCR buffer (50 mM Tris HCl, pH 8.3 @ 25°C, 50 mM KCl, 1.5 mM $MgCl_2$), 200 μM dNTPs, 200 pM with respect to each primer and 1.5 U of Taq DNA polymerase. A touchdown PCR protocol (Don, R.H. et al., *Nucleic Acids Res.* 19:4008 (1991)) was established to facilitate the robust amplification of most markers under the same conditions. PCR was carried out at 94°C for 45 sec., 52°C for 30 sec., and 72°C for 1 min.

The microsatellite markers were initially evaluated in ten sets of parents from the BT pedigrees. Those markers for which at least one parent was heterozygous were then evaluated in all the dogs in the pedigree. Seven to twelve microliters of product were run on a 5% to 7% Hydrolink D600 acrylamide horizontal gel according to the manufacturer's instructions with the following modification. During the overnight runs, a plexiglas gel carrier was placed on top of the gel to prevent the swelling and distortion that was otherwise observed. Initially, electrophoresis was carried out from 4 to 5 hr. at 50 V in 1 X TBE (90 mM Tris, pH 8.3, 90 mM boric acid, 2 mM EDTA) with ethidium bromide. A photograph was taken and the gel

electrophoresis then continued overnight at 35-40 volts depending on the fragment size of the product. A second photograph was taken and the results visually evaluated. It was found that two photographs were helpful in comparing different dogs with similar patterns. The alleles were then tabulated and used in linkage analysis.

Linkage analysis. Two point LOD (logarithm of odds) scores between CT and all the markers tested were generated using the MLINK program of the LINKAGE package (v5.1) (Lathrop, G.M. et al., *PNAS (USA)* 81:3443-3446 (1984)). A gene frequency of 0.5 was assumed for CT.

Results

Two hundred thirteen microsatellite markers were evaluated in the process of finding linkage. Of these 213 markers, 181 provided scorable products in BTs using the touchdown protocol described above. Of these, 114 were informative in the pedigrees and were further evaluated.

Of all the markers tested for linkage to CT, only one yielded a significant LOD score. As shown in Table 5 below, marker number C04107 was found to be linked to the CT locus at a LOD score of 5.96, at a recombination fraction of zero. No recombinants were detected. Since a LOD score of 5.96 indicates that the odds of observing this linkage by chance is about 1 in a million, and since, a LOD score of greater than 3 or an odds ratio of 1 in 1000 is considered proof of linkage, the findings imply that the CT locus is indeed very close to the C04107 locus and thus can be used to predict the inheritance of alleles at the CT locus. No recombinants were detected in this study and thus a value can not be put on the genetic distance between these loci, except to say that they are very close.

Table 5

θ (Recombination Fraction):	0.0	0.001	0.01	0.05	0.15	0.1	0.2	0.3
C04107 vs. CT	5.96	5.95	5.85	5.38	4.78	4.14	3.49	2.13
C04107 vs. ESD	-∞	-19.73	-10.78	-4.77	-2.44	-1.28	-0.6	-0.01
C04107 vs. RB1	-∞	-20.35	-11.43	-5.47	-3.18	-2.01	-1.28	-0.47

The primer sequence and allele information about this marker are shown in Table 6. The allele frequencies were determined from alleles observed in apparently unrelated dogs.

Table 6

Marker Locus	C04107
Repeat Motif in Reference Clone	$(CA)_6$ $CT(CA)_{11}$
Primer Pair	TCAGCAACTATACATTTAAGAGGA CTGTCCCATCTAAAGGATAGG
Allele 1 and Frequency	163 bp, 0.39
Allele 2 and Frequency	167 bp, 0.61

Marker C04107 was used to locate markers C04107B and C04107C shown in Table 2A, which are close to C04107 and also contain repeats. This "family" of markers may be used to detect CT.

A typical pedigree illustrating linkage to C04107 is shown in Figure 3. In Figure 3, circles and squares depict females and males, respectively, and individuals affected with CT are indicated by the filled symbols. The asterisk in the figure indicates an individual not available for analysis. The bands are the negative image of amplification products obtained from the dogs indicated in the pedigree and analyzed individuals share the 2,2 genotype at this locus. In this pedigree, all dogs with the 1,1 genotype are predicted to be homozygous normal while those with the 1,2 genotype are predicted to be heterozygous, and thus carriers of the CT gene.

Given the finding of linkage and allowing for a small error for recombination, it is predicted that all the offspring with the 1, 1 genotype are clear of the CT gene *i.e.*, homozygous normal, and that all 1, 2 offspring are carriers in this pedigree.

Since data on the ESD and RB1 loci were available for most of the dogs from a previous study (Yuzbasiyan-Gurkan, V. et al., *Genomics* 15:86-90 (1993)), the linkage relationships of these loci with C04107 were was also evaluated. Neither ESD or RB1 were found to closely linked to C04107 (see Table 5).

As demonstrated by the pedigree illustrated in Figure 3, given an informative mating, it is now possible to identify all the genotypes in the offspring, distinguishing between the homozygous normal, homozygous affected and heterozygous dogs provided the genotype of one affected dog is available. However, C04107 is not extremely polymorphic in the BT population, showing only two alleles and a

calculated heterozygosity of 0.43. Therefore, typing at the C04107 will not always yield information about the CT status of the offspring. Thus far, all affected dogs have been of the 2,2 genotype and the 2 allele is more common than the 1 allele (see Table 6). The matings which produce affected dogs will be found to be either between parents who are both 2,2 both 1,2 or one 1,2 and the other 2,2. In such cases, typing at the C04107 locus will only be useful in the second and third mating types. In the latter mating pairs, predictive information would only be available as to which dogs are affected. In order to make most pedigrees in the breed informative, additional polymorphic markers closely linked to C04107 are developed. It is predicted that a battery of three to five highly polymorphic markers will make almost every pedigree informative.

If strong linkage disequilibrium occurs at C04107 or nearby loci, the predictive power will be substantially improved. However, further studies of allele distributions in the BT population are needed to evaluate linkage disequilibrium. In any case, it should be possible to dramatically reduce the frequency of this serious disease within a very few generations.

As discussed above, canine copper toxicosis is present in the West Highland White Terrier and perhaps in several other breeds. (Thornburg, L.P. et al., *Vet. Pathol.* 27:81-88 (1990)). In the West Highland Terrier, it is clear that the phenotype is more complex, in that there is a spectrum of liver copper levels. This marker is evaluated in the West Highland White Terrier breed and it is determined whether there is segregation of high liver copper values with C04107.

The foregoing discussion discloses and describes merely exemplary embodiments of the present invention. One skilled in the art will readily recognize from such discussion and from the accompanying claims and drawings, that various changes, modifications and variations can be made therein without departing from the spirit and scope of the invention.

All publications referred to herein are expressly incorporated by reference.

WE CLAIM:

1. A primer comprising a polynucleotide, wherein the polynucleotide has a sequence selected from the group consisting of the sequences of Table 2A.

2. The primer of Claim 1, wherein the sequence is the Sns sequence of marker locus C04107 of Table 2A.

3. The primer of Claim 1, wherein the sequence is the Asn sequence of marker locus C04107 of Table 2A.

4. The primer of Claim 1, wherein the sequence is the Sns sequence of the marker locus C04107B of Table 2A.

5. The primer of Claim 1, wherein the sequence is the Asn sequence of the marker locus C04107B of Table 2A.

6. A method for amplifying DNA, comprising the step of performing PCR with the DNA and a primer set selected from the group consisting of the primer sets of Table 2A.

7. The method of Claim 6, wherein the primer set is that shown as the Sns sequence and Asn sequence of the marker locus C04107 of Table 2A.

8. The method of Claim 6, wherein the primer set is that shown as the Sns sequence and Asn sequence of the marker locus C04107B of Table 2A.

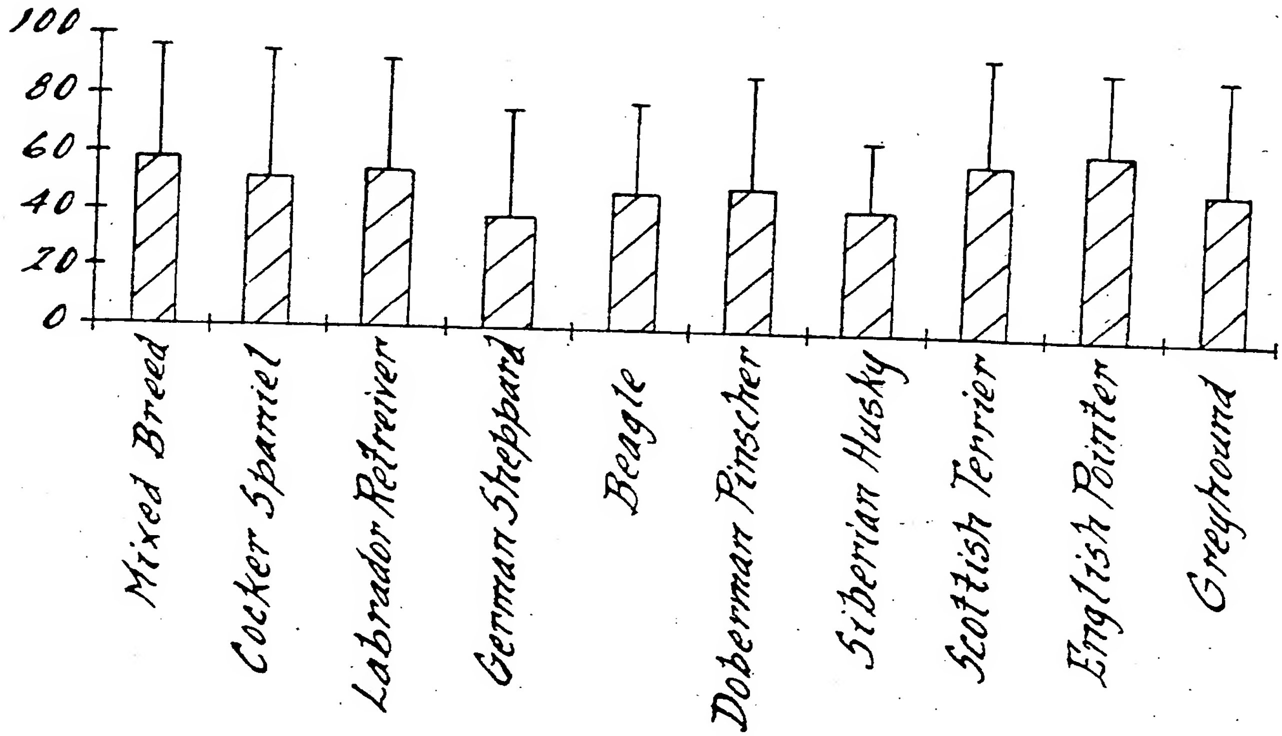

100
80
60
40
20
0
Mixed Breed
Cocker Spaniel
Labrador Retreiver
German Sheppard
Beagle
Doberman Pinscher
Siberian Husky
Scottish Terrier
English Pointer
Greyhound

FIG. 1A.

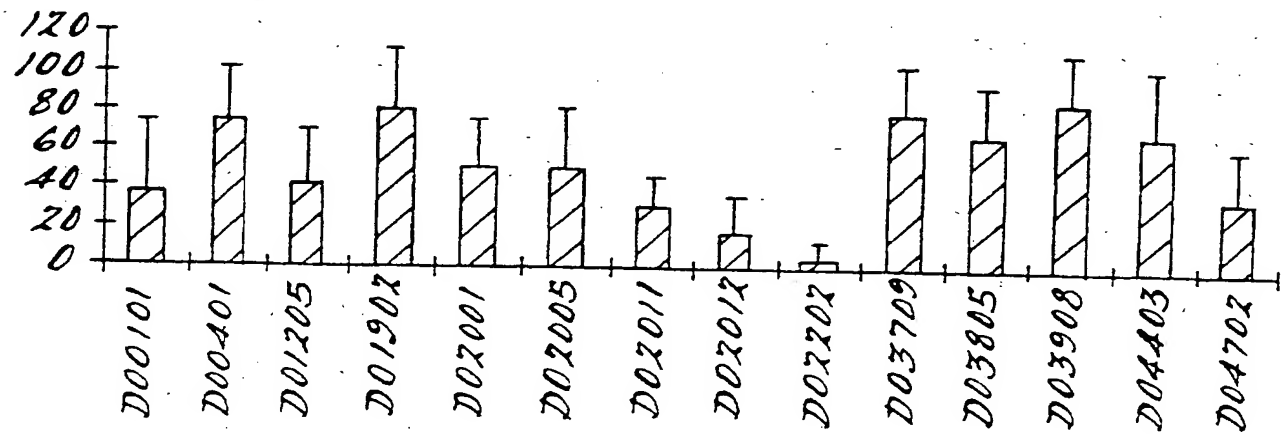

120
100
80
60
40
20
0
D00101
D00A01
D01205
D01907
D02001
D02005
D02011
D02012
D02707
D03709
D03805
D03908
D04403
D04702

FIG. 1B.

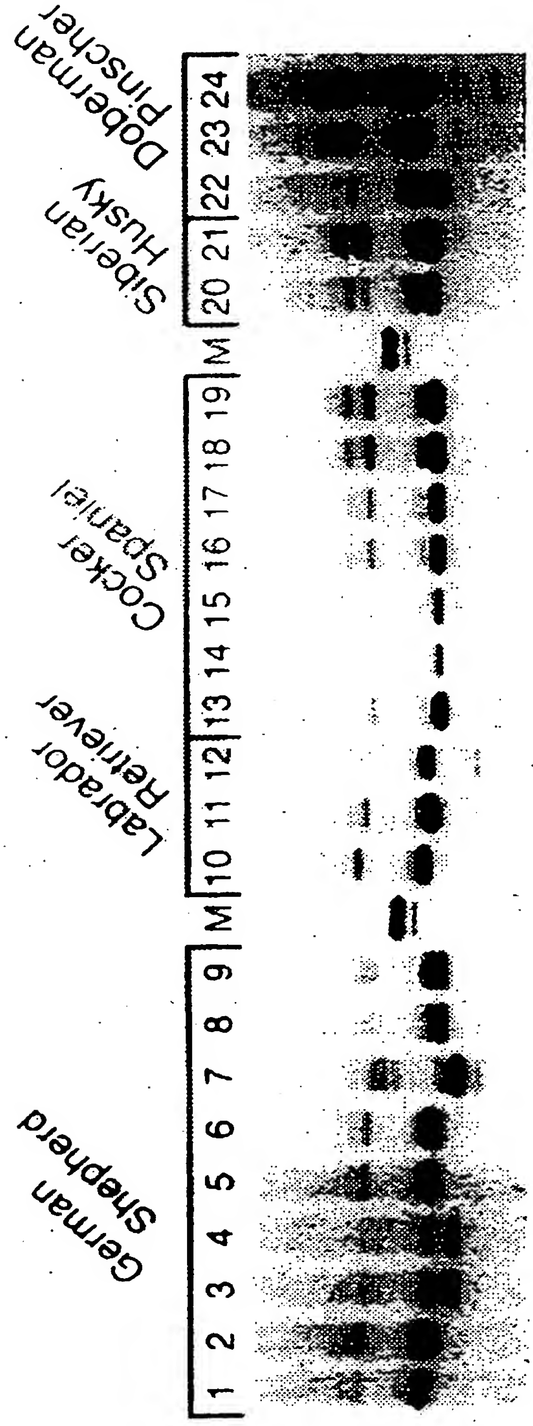

German Shepherd
1 2 3 4 5 6 7 8 9
M
Labrador Retriever
10 11 12 13
Cocker Spaniel
14 15 16 17 18 19
M
Siberian Husky
20 21 22
Doberman Pinscher
23 24

FIG. 2A.

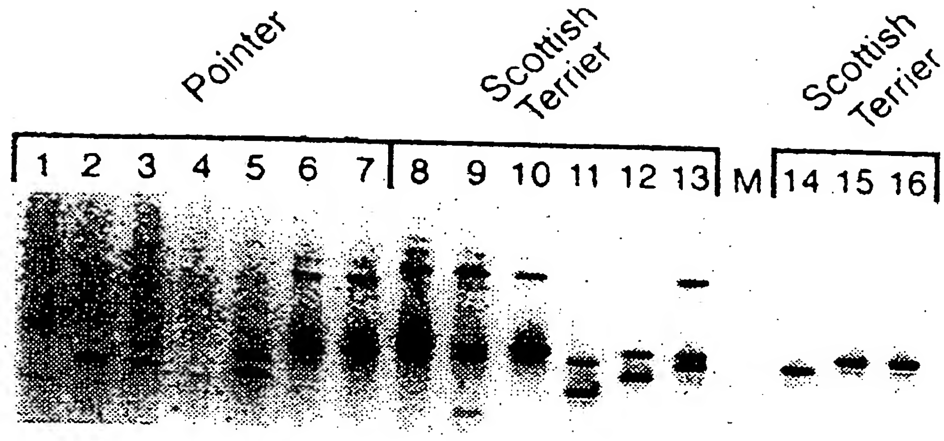


Fig. 2B.

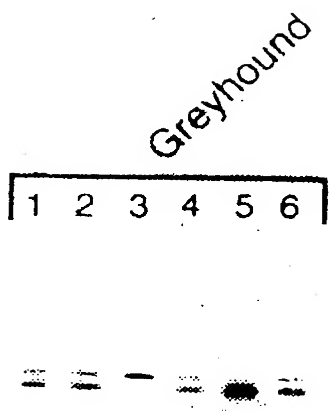


Fig. 2C.

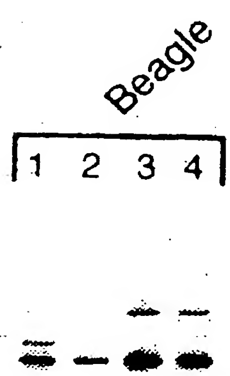


Fig. 2D.

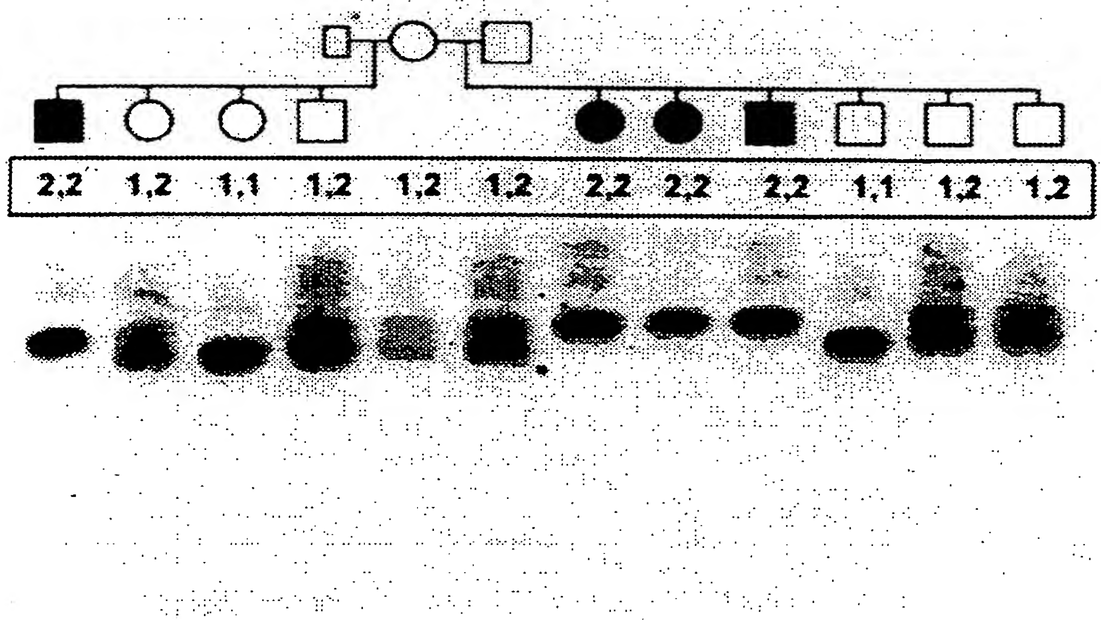


FIG. 3.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02396

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07H 21/04; C12Q 1/68
US CL : 536/24.33; 435/6

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1, 24.33; 435/6, 91.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	OSTRANDER et al. One hundred and one new simple sequence repeat-based markers for the canine genome. Mammalian Genome. March 1995. Vol. 6, No. 3, pages 192-195, especially abstract and Table 1.	1-8 (in part)
Y	OSTRANDER et al. Identification and Characterization of Dinucleotide Repeat (CA)n Markers for Genetic Mapping in Dog. Genomics. April 1993. Vol. 16, No. 1, pages 207-213, especially Table 2.	1-8 (in part)
A	YUZBASIYAN-GURKAN et al. Linkage Studies of the Esterase D and Retinoblastoma Genes to Canine Copper Toxicosis: A Model for Wilson Disease. Genomics. January 1993. Vol. 15, No. 1, pages 86-90, especially page 86.	1-8 (in part)

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search 10 JUNE 1997	Date of mailing of the international search report 08 JUL 1997
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DEBRA SHOEMAKER Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02396

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	FREDHOLM et al. Efficient resolution of parentage in dogs by amplification of microsatellites. Animal Genetics. February 1996. Vol. 27, No. 1, pages 19-23, especially page 21.	1-8 (in part)
A	ROTHUIZEN et al. The incidence of mini- and micro-satellite repetitive DNA in the canine genome. Theoretical and Applied Genetics. October 1994. Vol. 89, No. 4, pages 403-406, especially pages 405-406.	1-8 (in part)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02396

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-8, as limited to 10 sequences

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

searched for inventors and keywords:microsatellite or linkage or polymorphism or allele and dog/canine genome or gene or dna and ca repeat and copper toxicosis in APS, CAPLUS, MEDLINE, SCISEARCH, LIFESCI, EMBASE, BIOSIS WPIDS. Searched sequences of elected group by registry, genbank and dgene.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

each of the 519 microsatellite markers disclosed in Table 2A are distinct species. It is noted that in two cases there are more than one primer set corresponding to the same loci, for example C01407, C01407B and C01407C, which do have unity with each other.

The claims are deemed to correspond to the species listed above in the following manner:
Claims 1 and 6 are generic to each of the 519 microsatellite markers disclosed. Claims 2-5 & 7-8 have unity with each other because a single microsatellite locus is claimed but do not have unity with claims 1 & 6 because distinct microsatellite loci are claimed.

The following claims are generic: 1 & 6.

Applicant is allowed to select 10 sequence for the search fee and pay an additional $200 for each adiditonal 4 sequences to be examined. Since there is unity of invention between C01407, C04107B and C01407C, these sequences are considered to be one speices. A search report will be established on C01407, C01407B and C01407C and the first four primer pairs (so as to form a group of 10 sequences) recited in Table 2A if no other groups ar paid for and it considers that the International Application does not comply with the requirements of unity of invention (Rules 13.1, 13.2, 13.2) for the reasons indicated below:

The species listed above do not relate to a single inventive concept under PCR 13.1 because, under PCT Rule 13.2,the species lack the same or corresponding special technical feaures for the following reasons: each of the 519 microsatellite markers claimed in claims 1& 6 are drawn to a unique nucleic acid squence, each with a unique location in the canine genome and each linked with distinct genes and traits. Thus there is no special technical feature that relates to these microsatellite makers to each other.

THIS PAGE BLANK (USPTO)